

Strength in unity

Adapting to reunification has been a painful process for much of the scientific community in the former East Germany. But the city of Magdeburg has shown how a successful outcome can be achieved.

It is a truism that upbringing influences character. Nowhere is this more clearly demonstrated than in the research institutes in the former German Democratic Republic, where, in the period since reunification, west and east German scientists have been thrown together, often uncomfortably. During the restructuring of research organizations along western lines, for example, many senior east German academics lost their jobs to west Germans when most academic posts were opened to competition. Some were deemed unemployable because of their political past. Others were simply not felt to be at the cutting edge of research.

Five years after completion of this unique exercise in social engineering, however, some of the rewards are now emerging. Magdeburg is a shining example (see page 725). The quality of research now being undertaken, particularly in the neurosciences, is as high as anywhere in west Germany. East German scientists who kept their posts say their work is more interesting and fulfilling than before; west Germans appreciate the opportunity to build something new and significant, particularly at a time when support for research has contracted in the west. So good is the current situation that one senior researcher last year turned down a prestigious chair in Bonn to continue his research in Magdeburg.

But the undercurrents created by a strong and continuing east-west culture clash tell a different story. West German scientists are frustrated by the tendency towards a bureaucratic and cautious approach to problems of east Germans, brought up in a society where it was unwise to have strong opinions, or to make decisions without

the support of one's superiors. Appeals to eastern researchers from western colleagues to use their initiative are still sometimes seen by the former as potential traps.

East German scientists continue to resent what they feel to be colonization by westerners who tend to dismiss their entire pre-reunification lives as having little value, and western values that have ruined the careers of many former colleagues. They often view their new west German bosses as arrogant and aggressive in their demand for constant progress, and see an enforced continual competition for jobs and research money as unnecessarily stressful.

In practice, the two sides usually manage to work well together, largely because, whatever the underlying tensions, it is easier, more fun and more satisfying for a researcher from the east to work in a well-equipped modern laboratory, with much greater personal freedom than before. But, understandably, experienced researchers seldom choose to mix socially.

It is therefore reassuring to see how, only eight years after reunification, young scientists already appear free of the same tensions. Those who were still undergraduates in 1990 have been brought up under a single system of research training, and, moreover, appear able to enjoy one another's company both during and after working hours unburdened by recent history. This, together with the high calibre of the research staff appointed from the west — by no means the case in all parts of east Germany — has been an essential element in the success of Magdeburg in establishing itself as a centre of research excellence. It is also a hopeful sign for the future of German science. □

Science advice in an uncertain world

Despite pressures to the contrary, debate on the scientific dimension of policy decisions must take place openly.

Last week, a large agricultural products company whose trial plots of genetically modified crops had been attacked by protesters came up with a simple proposal: that in future, the precise location of experimental crops should be kept secret. The suggestion has obvious attractions. But it would also be a dangerous and misguided move. For nothing generates distrust of science and technology more than excessive secrecy, and the feeling that, if something is being kept out of the public eye, it must be potentially embarrassing — if not dangerous.

The same need for openness applies to science advice to governments. In an earlier era, science could count on public trust, and such advice could be drawn up as a matter of course by expert committees and submitted to decision-makers on a confidential basis. But today — a lesson learnt in the United States through the Watergate affair, and more recently in Britain by the BSE crisis — confidence in political decisions requires confidence in the political process. And this requires responsiveness to the electorate's demand for transparency.

The BSE crisis weighed heavily on the mind of Sir Robert May, Britain's chief scientific adviser, when he made an eloquent plea last week in defence of openness (see page 720). According to Sir Robert, this process should extend to the debates between scientific experts. If

respected advisers were encouraged to air their differences in public, he argued, everyone would get a better idea of how scientific knowledge is built up, not by maximizing certainty or striving awkwardly for consensus, but by minimizing uncertainty.

A different but equally radical proposal comes in a submission to the same inquiry from the Royal Society. This is the notion that the chief scientist should have access not only to the principal science advisers in all government departments, but also to debates on scientific issues taking place within those departments; again, BSE provides a demonstration of what can happen when such access is denied.

All such proposals, of course, will meet counter-arguments. Those concerning commercial confidentiality or national security are often (though not always) valid. Others, such as the tradition that policy-related questions require the type of clear, unambiguous answer that science is often unable to provide, can be less so. There will always be those who exploit openness for their own ends, as those responsible for destroying the experimental crops seem determined to do. And there are times when a clear statement of scientific consensus is essential. But the ability to handle openly the type of uncertainty that scientific knowledge can lead to is itself a form of social maturity. □



Journal prices lead libraries to back less costly initiatives

[WASHINGTON] A coalition of North American university libraries is planning to launch a bid next week to reduce the power of commercial scholarly publishers by endorsing new publishing ventures that it considers give value for money.

The initiative is being launched by the Scholarly Publishing and Academic Resources Coalition (SPARC), a body set up last October by the Association of Research Libraries (ARL). The first in a series of partnership agreements with journal publishers is due to be unveiled next Tuesday at the annual conference of the American Library Association in Washington, DC.

Under such agreements, publishers will undertake to launch high-quality, low-priced journals in fields that are currently dominated by expensive titles. In return, SPARC will guarantee a subscription base of research

libraries, and help to generate support for the new journals from faculty, learned societies and educational organizations.

The SPARC project reflects a growing concern among the coalition's members, which include prominent research universities such as Harvard University, the University of California at Los Angeles and the University of Toronto, about the rapidly increasing cost of library materials, especially research journals.

Last year, ARL estimated that, despite having cancelled hundreds of thousands of dollars' worth of subscriptions, research libraries were still spending 124 per cent more than they had in 1986, in order to acquire 7 per cent fewer titles. Between 1986 and 1996, the cost of scholarly journals increased by 148 per cent — far outpacing growth in both the consumer price index (44

per cent) and university operating budgets.

Mary Case, director of ARL's office of scholarly communication, says SPARC intends "to develop prestigious, cost-effective alternatives" to existing journals. "The aim is to lure faculty, publishers and editors away from high-priced titles," she says.

Eighty of ARL's 121 members are founding members of SPARC, and have each contributed \$5,000 to a start-up fund for the project. SPARC will not publish journals itself, but will enter into partnerships with publishers that share its aims.

According to Case, these encompass not just lower cost, but "access under terms that include educational uses and fair use — for example, giving [publication] rights back to the author". (Demand for the latter has increased significantly with the advent of personal Web sites and public 'e-print' servers.)

The first publisher to sign up with SPARC will be the American Chemical Society (ACS), which already publishes 26 journals in print and electronic form. ACS is planning to announce next week what it hopes will be the first of a number of new journals launched under the agreement with SPARC. "They can assist us to do what we want to do anyway — expand our line of publications," says ACS's publications director, Bob Bovenschulte.

SPARC's assistance in this case will be non-financial; according to Bovenschulte, they will merely endorse the new publication, and "help to sell subscriptions" — for example, by educating faculty and university administrators about the need to reduce the costs of scholarly communication. Looking ahead, Bovenschulte says that ACS hopes to launch at least one new journal a year in partnership with SPARC.

Bovenschulte believes that in general — and unlike commercial publishers — professional societies have not been as responsive as they might have been in launching new publications as new sub-fields of established disciplines have emerged.

"We've all ceded a lot of ground to commercial publishing houses who have been more aggressive and had a lot of investment capital," he says. "We haven't been starting new publications with risk capital."

Case emphasizes that SPARC is not seeking to compete with all commercial publishers, but only to challenge those it considers impose unfair profit margins or excessively restrictive copyright policies. "I envisage that we would [work jointly] with both commercial and non-commercial publishers," she says. "We're looking for partners who share our values."

Laura Garwin

'Insider' gets energy secretary nomination

[WASHINGTON] President Bill Clinton last week nominated Bill Richardson, the US ambassador to the United Nations and a former congressman from New Mexico, as the next energy secretary. The 50-year-old Richardson is a Clinton friend, a Democratic insider and, like outgoing energy secretary Federico Peña, a Hispanic.

He is also well acquainted with the department he is to head. From 1983 to 1997, he represented the congressional district that includes two key Department of Energy (DOE) labs, Sandia and Los Alamos, which he helped to protect from budget cuts as a member of the House commerce committee.

New Mexico's interests are likely to remain crucial to DOE politics if Richardson's appointment is confirmed. He is thought to have his eye on governorship of that state, or perhaps a role as Al Gore's running mate in the next presidential race. Pete Domenici, New Mexico's Republican Senator, already chairs the appropriations panel that determines the DOE's budget.

Richardson brings to his new job a strong interest in international affairs. Even before his move to the UN, Clinton used him in hostage negotiations with several countries, including Cuba, Iraq and North Korea. He has twice been nominated for the Nobel peace prize.

In a White House ceremony to announce his nomination, he mentioned the value of technical advances for environmental cleanup and energy research. He also spoke of the need to maintain a safe and reliable



Richardson: expected to be confirmed in post, but may face some embarrassing questions

nuclear stockpile to better support the Comprehensive Test Ban Treaty.

In Congress in the early 1990s, Richardson was sharply critical of DOE plans for the Waste Isolation Pilot Plant, the controversial underground nuclear waste repository in his home state, and he worked to ensure stringent environmental standards for the project. Ironically, he is likely to preside over the plant's early operations (see *Nature* 393, 199; 1998).

Richardson's nomination is not likely to meet resistance in the Senate, which has to confirm the appointment. But he may face embarrassing questions about his minor role in the scandal involving former White House intern Monica Lewinsky, whom he offered a job at the UN last year.

Tony Reichhardt

UK research set for international scrutiny

[LONDON] Panels of overseas scientists will for the first time help to assess the quality of Britain's best university research, according to Britain's higher-education funding councils. The move is one of several changes to Britain's Research Assessment Exercise (RAE), a regular initiative to assess the quality of university research. The next RAE will begin in April 2001.

"Science is an international activity," says Bahram Bekhradnia, director of policy at the Higher Education Funding Council for England (HEFCE). Funding councils, he adds, will now have the chance to calibrate Britain's research on an international scale.

Each of the RAE's 60 subject panels will choose a corresponding 'virtual' panel of scientists overseas. Members of the overseas panel will not meet physically, but will review all research considered by the UK panel to be worthy of the highest classification.

This is not the first time overseas scientists have been asked to help assess Britain's research; they have occasionally been co-opted on to UK-based panels before. But Bekhradnia says such an arrangement proved unworkable for logistical reasons; it was therefore decided to set up separate panels of overseas experts.

Other changes to the previous 1996 exercise include giving Britain's 117 universities longer to prepare their submissions, and giving feedback on performance. Panel chairs will be nominated by the entire outgoing panel, instead of just the outgoing chairman.

From next month, HEFCE will begin to consult universities on how to improve university teaching while retaining high-quality research departments. This is partly in response to the widespread perception among universities that the university funding formula encourages universities to invest in research at the expense of teaching.

The National Committee of Inquiry into Higher Education chaired by Lord Ron Dearing suggested a cash incentive to dissuade mediocre research institutions from diverting resources away from good teaching departments to boost research. But this idea, according to a HEFCE spokeswoman, was not well received.

The RAE is designed to help the higher-education funding councils ensure that their research funds of £826 million (US\$1.38 billion) are weighted so that larger amounts of money go to universities whose research departments are the most productive.

It works mainly through peer review, with 60 expert panels assessing 69 subject areas. Each university subject-group is given a quality rating on a seven-point scale, based on an assessment by the panel of each individual member's four best pieces of research.

Subject groups are scored from 1 — the

lowest rating — to 5*, which indicates that most of the work submitted is equivalent to the best in the world. A score of 3 is divided into two categories: 3b indicates that most research is of national excellence; 3a may also include research of the highest international standard.

Despite the planned changes, the announcement of the new RAE is likely to reopen many of the criticisms prompted by previous assessments. A major concern of the Association of University Teachers, for example, is that the RAE may jeopardize future research talent by cutting off funds from promising but less well-established researchers who are working in low-scoring departments.

But Bekhradnia disagrees, pointing out that money is withheld only from those universities whose departments score below 3b, and only for five years.

Other critics believe that the RAE encourages university departments to focus their expertise on areas of national priority — such as those identified by the Technology Foresight programme — to attract research

funding, when they should be setting their own research agendas.

This is a particular worry for the Engineering and Physical Sciences Research Council, which already spends around 70 per cent of its budget on Foresight-related activities. David Leech, the council's director of planning, believes that the research councils and the funding councils should maintain a distance from each other in the kinds of research they fund, to ensure diversity within British research.

Diana Garnham, general secretary of the Association of Medical Research Charities, and an RAE assessor, agrees. "It is important for universities to be able to initiate ideas for us [the charities] to respond to. We don't want to direct all of the research."

But again Bekhradnia says such fears are unfounded. "It is up to universities to decide how to spend money they receive from the funding councils," he says. "They may want to give it to their existing centres of excellence. On the other hand, they may want to develop other areas. These are decisions for them, not for us."

Ehsan Masood

Plea for plurality on advice to ministers

[LONDON] Britain's top scientific adviser has urged that the full range of scientific advice to governments should be made public — and not just a consensus view — in a bid to convey a sense of the different scientific opinions on issues.

Addressing the House of Commons Select Committee on Science and Technology last week, Sir Robert May, the prime minister's chief scientific adviser and head of the Office of Science and Technology, said that public confidence in the scientific advisory system needed to be regained following controversies such as the bovine spongiform encephalopathy (BSE) crisis.

Trust could be rebuilt if science advisers "consult widely, and consult contrary opinion" when issuing advice, said May. He added that all opinions should be made public, along with — wherever possible — advice to politicians based on them.

May acknowledged that civil servants were "uncomfortable" with the approach he was advocating, as it conflicted with the tradition that advice to ministers must remain confidential. But he said that, although there were genuine instances where confidentiality was paramount, "in general, I am for openness and taking risks".

May, who is Royal Society research professor at the University of Oxford and Imperial College, London, made his comments during a hearing at the House of Commons held as part of an inquiry into the



May: need to rebuild public confidence.

effectiveness of the scientific advisory system.

May said he was satisfied with the level of scientific awareness among government officials. But BSE provided one example in which research awareness among civil servants could have been better.

With the exception of the agriculture ministry, May said he considered government departments' understanding of relevant scientific issues "exemplary". By contrast, he said the government was less successful in its attempts to gain public confidence in the scientific advisory process.

Not all members of the select committee were convinced by May's remarks. Nigel Beard (Labour, Bexleyheath and Crayford) questioned the perceived benefits of publicizing conflicting views, particularly on issues relating to food safety. He suggested that this could lead to confusion among the public.

But May said that a commitment to openness and an acknowledgement of differing views had contributed to the success of the government's AIDS awareness campaigns during the late 1980s and early 1990s.

E.M.

Defence sanctions open doors for India's private companies

[NEW DELHI] The Indian government announced last week that it will allow private companies to become more closely involved in defence research and development, as well as in the production of military hardware.

The move is being widely interpreted as an attempt by the government to overcome the impact of the sanctions announced by the Clinton administration in the aftermath of the nuclear tests conducted by India last month (see *Nature* 393, 197; 1998).

According to defence minister George Fernandez, some of the nation's secret defence laboratories will be opened to the private sector, and private companies have been invited to become 'partners' in defence research. At present, the private sector has no role in defence research, and almost all of India's military hardware is produced in factories which are under the control of the defence ministry.

Fernandez announced the new policy at a meeting with leading industrial companies that are members of the confederation of Indian industry (CII). Rajesh Shah, CII's president, assured the government that the private sector would be able to play a big part in defence projects.

In spite of official claims that the US sanctions will not harm India's defence efforts, defence scientists privately admit that most of the projects run by the Defence Research Development Organization (DRDO) involve many imported components to which India may no longer have access.

It is also feared that the sanctions may hamper DRDO's plan to increase the indigenous content of military hardware to 70 per cent by the year 2005, from its current 30 per cent. The only way of achieving this goal now is by increasing the involvement of Indian industry.

Besides working with DRDO as partners in research, private companies will, under the new policy, be allowed access to technology developed by DRDO with potential civilian use. According to DRDO chief Abdul Kalam, for example, techniques for making carbon composite material used in missile nose cones will be available to a company that wants to produce this material to make lightweight calipers for polio victims.

A special committee will choose the DRDO technologies with potential civilian applications. Other committees will identify research areas for partnership between DRDO and CII companies, and possible joint ventures for producing and exporting military hardware.

K. S. Jayaraman

Russian miners add weight to protests by scientists

[MOSCOW] Researchers from branches of the Russian Academy of Sciences throughout the country held a mass protest march in Moscow last week to complain about the continued deterioration in their working conditions.

The protest followed the news that, after an earlier decision by the government to reduce funding for science next year by 27 per cent, it has now been decided that science will get only two-thirds of this reduced financing because of the country's worsening economic situation.

"Our leaders do not seem to understand that Russia cannot afford to allow its science to be destroyed, and should pass the steering-wheel to others," says Viktor Kalinushkin, a physicist and chairman of the Moscow council of the Academy of Sciences' trade union.

The protest march brought together representatives of all the scientific centres near Moscow, as well as others from 20 research centres throughout the country. It began on 15 June in unusually high early summer temperatures at the Pushchino biological centre, 100 kilometres from Moscow.

During the march, one group of scientists briefly blocked the Simferopol highway leading to the capital from the south-west, while another group blocked the Kaluzhskoe highway near Troitsk scientific centre, 30 kilometres from Moscow.

The scientists reached the Institute of Medicinal Plants in the suburbs of Moscow three days after setting out, and held a short meeting before boarding a bus to the city centre.

"Most Russian scientists consider the president and the government unable to alter the present political course, which turns Russia into a backward country of colonial-type economy," said Kalinushkin in a television interview.

"If they are just going to sell off the country's natural resources, then they do not need

high technologies, to say nothing of fundamental science. This explains why President Boris Yeltsin and his newly appointed prime minister Sergey Kirichenko are breaking their promises to scientists so easily."

Originally, the trade union had planned to hold a meeting at Gagarin Square, near the academy's presidium building, with 3,000 participants. But the city authorities refused permission, arguing that it would interfere with traffic.

After the scientists had rejected alternative sites, it was suggested that they should hold their meeting near parliament, provided there were no objections from the miners, who had been holding a protest there for several days.

The government may have been hoping that the miners would show little desire to cooperate with the scientists, whose previous complaints have been purely economic. But the two groups found themselves united in demands to change the regime, and the miners agreed to a joint protest.

A decision by the authorities to reduce the number of those allowed to take part in the protest to a few hundreds meant that the scientists were restricted to five or six representatives from each institute.

Vladimir Strakhov, director of the Earth Physics Institute and full member of the academy, was the only scientist of such high rank to take part in the protest. Eighteen months ago, he went on hunger strikes to protest against scientists not being paid (see *Nature* 383, 563; 1996).

The joint action of the miners and scientists provided a unique demonstration of the Marxist commitment to "the union of physical and mental labour". For the first time in Russian history, the two groups seemed united behind slogans that were virtually identical — which could lead to some of the most dangerous protests since the beginning of 'perestroika'.

Carl Levitin



Fraternity: scientists protesting over further cuts in research budgets greet striking miners at a protest outside the Russian White House. Placards read "Castration of science means impotence for Russia."

Sweden must close centres or leave CERN, says council

[MUNICH] Sweden's Natural Sciences Research Council (NFR) has warned the government that, unless budget cuts are restored, it will either have to close some national research centres or give up its membership of CERN, the European Laboratory for Particle Physics.

In an interim report delivered to the government last week, the NFR says it would recommend closing "two or three" of the national facilities, rather than leaving CERN or reducing university grant money, which has been heavily pruned in recent years.

The two most vulnerable facilities are the The Svedberg Laboratory, a particle physics laboratory in Uppsala, and the Onsala Space Observatory near Gothenburg, which has two radiotelescopes.

Last year the Swedish government imposed a 14 per cent budget cut on the NFR, which funds university research in biology, physics, chemistry and mathematics, as well as four national research laboratories. Threats of wider cuts in the research budget then prompted government officials to consider pulling out of CERN (see *Nature* 386, 636; 1997). But the research minister Carl Tham was persuaded not to do so.

Now Sweden's particle physicists face the same threat again. Earlier this year the government reorganized its budget lines to give the NFR responsibility for funding international organizations such as CERN and the European Southern Observatory (ESO). The government also imposed a cut in the annual budget of SKr42 million (US\$5.3 million), approximately equal to Sweden's annual subscription to CERN.

The NFR says that continued membership of international research organizations should be Sweden's highest priority if, as looks certain, cuts are not reversed. It is now trying to reach a consensus with universities on the best alternatives for handling the cuts, which come into effect in 2000.

The NFR cannot decide on its own whether or not to leave CERN. This decision lies with the government, which has frequently complained about the high cost of subscriptions to international research organizations. The Swedish particle physics and astronomy communities are served at CERN and ESO, although activities at national and international facilities do not overlap.

The NFR will present its final report to the government in the autumn, when a report from a parliamentary commission on the overall organization of research in Sweden is also due. If it were to leave CERN by 2000, the government would have to announce this by December. **Alison Abbott**

Congress remains upbeat on public genome efforts

[WASHINGTON] The chairman of a key congressional committee promised last week that funding for the Human Genome Project was not in jeopardy in the light of a recently announced private plan to sequence the entire human genome in only three years.

"I doubt very much that Congress will cut funding" for the project, said Ken Calvert (Republican, California), speaking to officials of the National Institutes of Health (NIH) and Department of Energy (DOE) during a hearing of his House subcommittee on energy and environment, which oversees DOE research.

Witnesses at the hearing generally viewed the joint venture of the Institute for Genomic Research (TIGR) in Rockville, Maryland, and Perkin-Elmer as complementary to the federal project, rather than likely to replace it (see *Nature* 393, 101 & 201; 1998).

Craig Venter, TIGR's president, tried to dispel notions of a race, saying his announcement "has led some to speculate that federal funding for the human genome is no longer needed. Nothing could be further from the truth." He called for more money to be spent on government genome research.

Francis Collins and Aristides Patrinos, who head genome research at NIH and DOE, respectively, tried to show how the Human Genome Project differed from Venter's plan. "The genome project is much broader than just the human sequence," said Collins.

The federal effort includes not just delivering raw sequence data, but also producing a high-quality database, distributing information to the scientific community, mapping genomes of other organisms, and studying the ethical, legal and social implications.

Collins said a five-year plan being developed for the federal project will use information from Venter's group (see *Nature* 393, 399; 1998). But he admitted it will be difficult for researchers to use the private data until their quality is demonstrated, which may take two or three years.

"The federal effort is fully prepared to adjust its strategy," Collins told the committee. But he added: "We should not drastically alter our strategy until we have more data."

David Galas, who headed DOE's part of the project before becoming president of Chiroscience R&D in Bothell, Washington, asked for researchers to publish a "first draft" of the human genome "as quickly as possible, whether or not the [Venter] effort contributes... to reaching this goal".

Galas estimated that speeding up publication of a draft version by even a year would save the private sector about \$2 billion in research costs.

The main dissent came from Maynard Olson, the University of Washington genetics researcher who chairs the NIH review committee for genome research. He predicted that there would be more than 100,000 major gaps in Venter's final assembled sequence.

But the committee members were clearly excited by the TIGR venture. Tim Roemer of Indiana, the senior Democrat on the panel, agreed it was important to be sceptical of promises that "appear too good to be true". But he said that Venter's plan held out the "golden possibility of a private-public partnership that could result in a phenomenal return for science". **Tony Reichhardt**

Patent office struggles to handle backlog

[MUNICH] The European Patent Office (EPO) is to recruit an extra 250 patent examiners next year in a bid to keep up with the rapidly growing number of applications.

The number of patent applications filed at the EPO rose by 14 per cent last year to a record 99,800, partly as a result of lower patent fees. And a backlog of 20,000 patent files is already clogging the in-trays of EPO patent examiners in Munich, The Hague, Berlin and Vienna.

Germany, Great Britain, Switzerland, France and the Netherlands submitted the highest numbers of patent applications last year, and European countries now account for more than half of all applications.

But, according to figures released at a press conference in Munich last week, all of the EPO's 19 member states — as well as the

United States and Japan — have increased their patent activities since the EPO agreed to reduce its official fees two years ago (see *Nature* 384, 506; 1996).

The largest number of applications was in the category of medical and veterinary technologies, which saw a rise of nearly 15 per cent from 1996, to 6,200. Electronic communications technologies (up 28 per cent) and biochemistry and genetic engineering (up 22 per cent) showed the highest rates of increase.

Ingo Kober, the president of the EPO, told the press conference that the costs of patent applications — which currently average DM60,000 (US\$33,500) — will be reduced further by lowering search costs, a move the EPO hopes will encourage innovation. **Quirin Schiermeier**

CNRS reform will boost strategic advice

[PARIS] The administrative council of France's Centre National de la Recherche Scientifique (CNRS), Europe's largest fundamental research agency, meets today (25 June) to suggest reforms that would reinforce the input of independent advice on scientific strategy at every level of CNRS management.

One key issue is the large role of the CNRS's national committee in setting overall research strategy. The committee is a significant body in French science, evaluating laboratories and administering recruitment not only within CNRS's own laboratories but also in CNRS laboratories attached to universities. Reforms to the committee will affect 80,000 staff.

The national committee has some 40 sections, each consisting of scientists from the research agencies and universities. According to one CNRS administrative council member, individual sections tend to support their own subdisciplines, rather than adopting a broader view of strategic needs.

"If you ask the laser physics community, of course they will say we need more research on lasers in France, but what we need to know is whether the entire physics community thinks this is the case," he says.

To remedy this, the council is likely to propose that, while the number of sections should not be drastically reduced for the purpose of day-to-day evaluation, there should be a major regrouping of sections when this is required to give a broader view of strategic issues.

Earlier proposed reforms of the national committee have met with resistance from researchers. A recent survey of all CNRS researchers by the administrative council showed that a large majority opposed a proposal by Claude Allègre, the minister for national education, research and technology, to halve the number of sections (see *Nature* 393, 506; 1998).

But Catherine Brèchignac, the CNRS director-general, is one of several administrators who believe the conflict over the number of sections has distracted attention from the more pressing issue of determining the committee's role in setting scientific strategy. She feels that the furore over the issue has triggered resistance among researchers that threatens to block more important reforms.

Another proposed change intended to reduce the power of individual sections is the complete remodelling of the councils that advise the CNRS's seven departments. At present these are made up of representatives of the sections of the national committee.

The council is expected to propose that each department should have an independent scientific board consisting of directly

elected researchers, nominated members and foreign scientists.

The recent consultation showed that a narrow majority of CNRS scientists — 54 per cent — agreed that the departments should play a larger role in setting strategy. A further 30 per cent who opposed this said they would accept it if assurances were given that the departments would base their strategic thinking on that of the national committee.

CNRS researchers' attachment to the national committee stems from the fact that two-thirds of its members are elected directly by the scientific community. But, as one council member points out, there seems no reason why researchers should not also accept separate departmental advisory committees, given that these would also include a large proportion of directly elected researchers.

Prospects for the proposed reforms seem to be improved by the fact that their broad thrust is supported by both Brèchignac and Edouard Brezin, the president of the administrative council, who have often clashed in the past over the content and pace of reforms.

The director-general's office is likely to see similar reforms. It has an advisory board of three ministry representatives and the

heads of the CNRS's seven departments, and is chaired by the director-general. Many believe this is an incestuous set-up and that a separate body, again including a significant number of international experts, is needed.

"The danger of the current set-up is that CNRS could be seen as just made up of seven independent departments, whereas it is supposed to take strategic decisions of national significance," says one CNRS official.

Under the reform proposals, the council itself would have a much greater role, being involved from the outset in preparing reforms and decisions. Although the board votes on the budget and other strategic issues, at present its role is largely restricted to rubber-stamping proposals from the director-general.

"This creates a situation where if, say, a proposal for genome strategy reaches us and we are unhappy with it, at present we only have two choices: accept it, or reject it, and provoke a crisis by setting back what has often been protracted negotiations among the various partners," says a CNRS council member. "We need to be involved from the outset to ensure that the director-general and the council have a parallel vision of the directions being taken."

Declan Butler

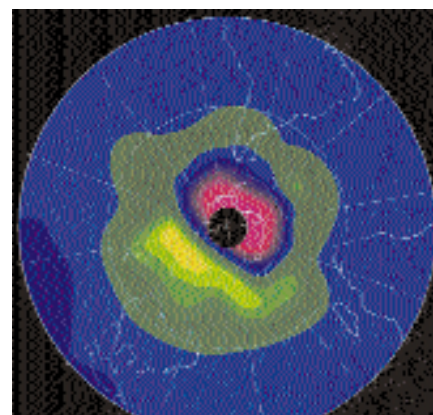
Ozone recovery will be long-term affair

[LONDON] Although the full recovery of the Earth's ozone shield is expected to occur by the middle of the next century, signs of the recovery may not become apparent for another two decades. That is the conclusion of the latest four-yearly ozone assessment published this week by the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP).

Godwin Obasi, secretary-general of the WMO, said the report, prepared by more than 200 scientists from around the world, showed that the 1987 Montreal protocol to phase out ozone-depleting substances is clearly working.

"The WMO global network of stations is detecting lower rates of increase in bromine and a decline in chlorine concentrations in the troposphere," Obasi said. But he added: "It might not be possible to detect firm signs of ozone recovery before another 20 years due to natural atmospheric and ozone variability."

Klaus Töpfer, executive director of UNEP, said the world owed scientists "a debt of gratitude" for their "unbiased advice", and for showing governments an "effective path to save the ozone layer". But he added that a complete recovery of ozone levels will need complete compliance with the



Still growing: the ozone hole is as big as ever in this view of the South Pole in October 1997.

protocol, particularly from developing countries.

Under the terms of the Montreal protocol, these countries must complete the phasing out of ozone-depleting chemicals by 2005. They are due to start next year. But emissions from India and China have exceeded earlier projections.

Developing countries have access to a global fund — the Montreal Multilateral Fund — to help them make the transition to using and producing alternatives to ozone-depleting chemicals.

Ehsan Masood

Attack on Japan's engineering education

[TOKYO] Japanese universities urgently need to introduce a system for evaluating engineering and technology departments if they are to produce internationally competitive engineers, according to leading industrialists and engineers in Japan.

Japanese companies say the quality of education in university engineering departments is often unclear, and that there should be a system for evaluating the standard of engineering graduates.

Takeshi Owa, chief fellow at Toshiba Research and Development Centre, says universities and companies have different objectives for education in engineering. "While industry wants highly trained engineers with technical skills, universities are aiming to create researchers," he says. "As a result, we are not entirely satisfied with the content of engineering programmes at universities."

Many are keen to see a body created to judge the quality of university curricula similar to the American Accreditation Board for Engineering and Technology (ABET) in the United States. However, such a move would be difficult to implement given the general reluctance of Japanese universities to endorse external evaluation.

An accreditation body would employ specialists to assess factors in university

departments such as the quality of teaching and the academic levels attained, as well as facilities and equipment. The aim would be to set a standard of education and training for engineers that would be a clear indicator to those seeking to employ engineers.

In the United States, where most universities have been evaluated by ABET, companies are required to employ engineers from ABET-approved universities in order to participate in public works projects.

Last year, the Ministry of International Trade and Industry (MITI) and the Japan Federation of Economic Organizations (Keidanren) joined the Japanese Society for Engineering Education and the Japan Federation of Engineering Societies to study ways of improving the standard of university engineering and technology departments.

"We are studying the system employed by accreditation bodies in the West, such as ABET in the United States, as a possible model for a similar organization in Japan," says MITI's Hideo Suzuki. According to Suzuki, MITI is keen to make such a body a significant part of its programme for improving Japan's ailing economy. "What we need most to rebuild our economy are people who are capable of making technical contributions to create new industries."

Officials at the Ministry of Education, Science, Sports and Culture (Monbusho), whose council on university education is compiling a report on an external evaluation system, say they support the idea of an independent assessment body for engineering. At the same time, they emphasize the importance of creating an evaluation system covering all university departments.

But engineers feel that a more rigorous step is needed to produce engineers of international standard. Their feelings are especially strong because the Asia Pacific Economic Cooperation forum is discussing the possibility of introducing an international qualification for engineers.

Itsuo Onaka, a professor of engineering at Osaka University, admits that the quality of engineers in Japan is poor.

"Universities seem to lack a clear objective for education in engineering and technology, and this makes engineering courses unappealing for many students," says Onaka.

He says the adoption of the system proposed by MITI would stimulate universities to improve their standards, and this would be particularly important if there were eventually to be an international qualification for engineers.

Asako Saegusa

South Africa's truth commission reveals bioweapons plot

[CAPE TOWN] South African researchers were involved in a sinister but bizarre effort to develop biological and chemical weapons under the previous government, according to evidence presented during the past two weeks at hearings of the country's Truth and Reconciliation Commission.

The research effort was carried out by the South African Defence Force. Some of its goals remain obscure. Johan Koekemoer, for example, former professor of organic chemistry at the Rand Afrikaans University in Johannesburg, said he was bewildered about plans to use a quantity of the drug Ecstasy worth one billion rand (US\$185 million).

The drug was manufactured in 1992, two years after the release of President Nelson Mandela and the lifting of the ban on the African National Congress (ANC). Some of it, along with Mandrax tablets, was used in experiments to create drug-laced tear gas. But the intended effect on rioters is unclear.

Uncertainty remains about what happened to the vast quantities of stockpiled drugs. Former surgeon-general of the South African Defence Force, Niel Knobel, giving evidence before the commission last week, said that hundreds of kilograms of chemical agents, including Mandrax, were dumped



off the southern Cape coast by the South African Air Force in early 1993. But Zenzile Khoisan, an investigator for the commission, suggested they might have been supplied, through gangster connections, to communities that were active in the struggle against apartheid.

The former head and chief architect of the weapons programme, Wouter Basson, faces multiple criminal charges, including instigating murder and the possession of drugs. Basson was due to appear before the commission last week, but has applied to the High Court to have his subpoena deferred

on the grounds that it would prejudice his criminal trial.

Basson was retired early from the South African Defence Force by President F. W. de Klerk in 1992, after an investigation into alleged illegal activities of the defence establishment — only to be re-employed by the Minister of Defence, Joe Modise, after Mandela's government was elected in 1994.

Basson liaised with researchers through a front company, Roodeplaat Research Laboratories, headed by former University of Pretoria veterinarian Daan Goosen, who became involved in the programme after being asked to supply snake venom intended to kill ANC members. Goosen was briefed to develop research projects including the selective poisoning of people on the basis of their skin pigmentation, producing a vaccine to reduce black fertility, and cultivating cholera and anthrax organisms.

It is unclear which of the programme's products were effective and which remained in the realms of morbid fantasy. Poison manufactured at Roodeplaat was used in an unsuccessful assassination attempt on the Minister of Justice, Dullah Omar. But anthrax spores were planted in the food of three Russian advisers to the ANC, one of whom subsequently died.

Michael Cherry

New voices displace east/west tensions

Alison Abbott

The east German city of Magdeburg, long known for its neuroscience research, went through a difficult period in the years after the Berlin Wall came down. But its scientific efforts received a boost last week with the opening of the Centre for Neuroscience Innovation and Technology.

[MAGDEBURG] Magdeburg, the east Germany city now famous for its high level of unemployment, last week celebrated the opening of a technology transfer centre that marks the success of post-reunification efforts to build a strong base in the neurosciences.

The Centre for Neuroscience Innovation and Technology (ZENIT) is on the university campus of a city that is still coming to terms with the cultural shock of the country's reunification. It was described by Hannelore Kohl, wife of the German chancellor, who opened the centre, as "an excellent example of how the scientific infrastructure has been successfully established in the new German *Länder* [states]."

ZENIT was built with the support of European Union structural funds, intended to stimulate the economies of poorer regions of Europe. It will complement a group of neuroscience research departments in the Institute for Neurobiology (IfN) and the Otto von Guericke University of Magdeburg.

Half of ZENIT's 5,000 square metres will be occupied by university neuroscience and medical technology departments. Five start-up companies spun off from the university and IfN are moving into the other half, where they will be joined by the new Max Planck Institute for the Dynamics of Complex Sys-

tems, and a state-of-the-art magnetic resonance imaging machine.

Magdeburg established a tradition in neurosciences during the communist era. Its Institute for Brain Research, the last institute to be founded (in 1985) by the East German Academy of Sciences, had an international reputation. In 1990 the Wissenschaftsrat, the science council of the federal republic, recommended that it continue to operate.

Two years later, a second Wissenschaftsrat committee recommended that the new university should focus its efforts on neuroscience. But the years immediately after the *Wende* — the political changes following the fall of the Berlin Wall in 1989 — have been difficult for east German academics.

Well over half of the professors lost their jobs. Some were suspected of having links with the Stasi, the East German secret police. Others were unwilling or unable to reapply successfully for their posts, as the government required.

In 1992, the Institute for Brain Research was reconstituted as the IfN, and re-hired only one of its seven research directors, Klaus Reymann. A year later the university was founded, merging Magdeburg's prestigious Medical Academy with other higher education institutes.

Of 15 new university chairs created in the neurosciences, most were filled by west Germans. Many professors left the Medical Academy to set up in private practice before their chairs were advertised as university positions, and others had their academic positions downgraded.

Five years later, the east German academics who survived this upheaval say they prefer the new, well-financed system, enjoying the freedom to travel and the efficiency of well-equipped laboratories.

But resentment lingers among many of those who lost their posts. Wolfram Neumann, an east German who is now dean of medicine at the university, says that, at first, doctors whose links with the university were cut refused to refer patients to university clinics. "But things are much better now."

The west Germans who moved to Magdeburg are also getting used to the poorer living conditions (now starting to improve) and an unfamiliar culture, factors that initially put off many scientists from moving east.

One who was not put off is Henning Scheich, who, at the age of 50, left a safe chair at the University of Darmstadt to head the IfN. He saw a unique opportunity to create a systems-based neuroscience institute with research activities at molecular, behavioural and clinical levels. His first "very unpleasant" task was to oversee the dismissal of many staff. "It was also difficult to encourage those scientists who remained to be creative," he says.

To "change the spirit", he brought in a small army of "young, self-confident west German graduate students". The IfN's 30 or 40 PhD positions are now filled by students from both west and east. Unlike their older colleagues, these young scientists are at ease working and socializing with each other.

For the older generation, the issue of west versus east will never disappear. Three years ago, for example, Reymann, an internationally acknowledged expert in the molecular mechanisms of learning, was dismissed from his post as an IfN research director after the Sachsen-Anhalt Ministry of Culture discovered a file allegedly linking him with activities helpful to the Stasi. "I started to travel to the west in 1986, and so I inevitably had to have contact with the Stasi," says Reymann.

Scheich ensured that the IfN kept Reymann on as a researcher using non-government grants. Reymann has now formed a neurosciences service company with other colleagues which is moving into ZENIT.

In doing so, Reymann joins an effort in Magdeburg — and in the whole of Germany — to harvest the results of academic research for industrial and social needs. Ironically, says Bernhard Sabel, head of the university's Institute for Medical Psychology and chairman of ZENIT's scientific council, this is something that universities in the communist east were much better at than their west German counterparts. □



The University of Magdeburg was named after Otto von Guericke, inventor of the air pump, who in 1658 demonstrated the force of atmospheric pressure. In an experiment that is regularly reconstructed (above), he used his air pump to create a vacuum between two hemispheres which could then not be pulled apart by 16 cart-horses. The hemispheres are incorporated into the logos of both the university and the Institute for Neurobiology, symbolizing to some the unification of West and East Germany.

Germany's grants chief wants more freedom for young researchers

[MUNICH] Ernst-Ludwig Winnacker, president of the Deutsche Forschungsgemeinschaft, Germany's main research grants agency, warned last week that talented young scientists could start rejecting university careers because strongly hierarchical structures prevent them from doing research independent of their department's professor.

Young researchers should be given responsibility for their own research much earlier than at present, Winnacker told the agency's annual meeting in Bonn. He also proposed abolishing the postdoctoral qualification *Habilitation* which is commonly required in universities in German-speaking countries for academic teaching posts. He believes it reinforces the academic hierarchy by tying researchers for too long to their professors.

Call for tough curbs on human cloning in Japan

[TOKYO] Research involving human cloning should be strictly regulated in Japan, according to an interim report released last week by the Council for Science and Technology (CST), the country's main science policy-making body. The report, compiled by the council's cloning subcommittee, calls for an effective control on human cloning by setting national guidelines similar to those on gene therapy, or a legal ban with penalties.

The separate Science Council, an advisory panel of the education ministry, released a report in February calling for a ban on research into human cloning at Japanese universities. But the CST proposes to impose regulations on all scientific institutions. It will seek the views of the public and scientific groups before making its final decision in the autumn.

Big three seek coherent space policy for Europe

[MUNICH] France, Germany and Italy last week called for a European Space Conference to encourage a more coherent space policy by bringing together all organizations involved, including the European Space Agency (ESA), the European Union, national space agencies, industry, research and user organizations.

In a joint statement, the three governments said that space activities should be more cost-efficient and applications-orientated. They said they "encourage" ESA's director general to

"continue his efforts" in this direction. But the statement is likely to irritate other ESA member states, which agree with the general sentiments but feel uneasy when ESA's largest members ally themselves publicly as a single block.

Wüthrich wins prize for research on proteins

[LONDON] A Swiss structural biologist, a Japanese mathematician and a Korean-born artist have won this year's Kyoto prizes. Kurt Wüthrich, professor of molecular biophysics at the Swiss Federal Institute of Technology, was honoured for his work in using nuclear magnetic resonance to determine the structure of proteins and nucleic acids (see *Nature* 382, 180–182; 1996).

Kiyosi Itô, professor emeritus of mathematics at Kyoto University, won the basic science prize, and the artist Nam June Paik won in the creative arts and moral sciences category. Each winner receives ¥50 million (US\$373,000). The prizes will be awarded in Kyoto in November.

Genetics field trials 'must be done in secret'

[LONDON] AgrEvo, the German agrochemicals company, is calling on the British government to conceal the location of sites used to test genetically modified crops. The move follows a spate of attacks on test sites in England and Scotland over the past two months.

A spokesman for AgrEvo said that five of the company's 40 UK test sites had been damaged. Attackers can easily locate test sites from a public register available on the Internet. The spokesman said: "The protesters are destroying valuable scientific data, which would help to answer many of the questions they themselves properly raise."

Gingrich puts science second only to drugs war

[WASHINGTON] Newt Gingrich (Republican, Georgia), speaker of the US House of Representatives and one of the most influential members of Congress, has recently been speaking out in support of science. In newspaper interviews, college addresses, and speeches at trade shows, Gingrich has sung the praises of government-funded basic research, and has called for a doubling of the federal investment in science over the next eight years.

Last week he told the *Washington Post* that science spending "ought to be our second highest priority after winning the war on drugs". Legislation to double the federal research budget, although over a longer timescale that some scientists had hoped, is expected to be introduced this month.

Congressman calls for inquiry on EPA officials

[WASHINGTON] The chairman of the House of Representatives science committee has asked congressional investigators to look into allegations that officials at the Environmental Protection Agency (EPA) have retaliated against scientists who expose fraud or challenge agency policies.

In requesting the General Accounting Office investigation, James Sensenbrenner (Republican, Wisconsin) cited a 10 June letter to the *Washington Times* signed by several EPA employees, including researchers at the agency's field centres. The letter alleges "egregious misconduct" by top agency officials. Sensenbrenner called the allegations "quite serious" and said they "have to do with the nature and quality of the science being performed at EPA". The investigation is expected to take at least six weeks.

Parliament set to probe troubled British Biotech

[LONDON] A UK parliamentary committee is to open an inquiry next month into British Biotech, the country's flagship biotechnology company, which is mired in controversy following alleged over-optimistic claims about the prospects for some of its drugs trials (see *Nature* 392, 852; 1998).

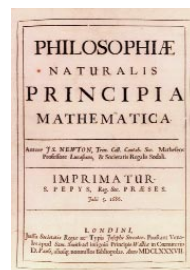
The timing of the inquiry is understood to have been accelerated to ensure access to information that may become limited following the news that Andrew Millar, the sacked former director of clinical research, and his former employers are to take legal action against each other. Millar is taking British Biotech to an industrial tribunal claiming unfair dismissal. The company is suing him for breach of confidentiality.

Newton's first edition goes under the hammer

[LONDON] A first edition of Sir Isaac Newton's *Philosophiæ Naturalis Principia Mathematica* (pictured) — mathematical principles of natural philosophy — was sold last week for \$321,000 at an auction at Christie's in New York. The amount is almost three times the estimated pre-auction value of \$120,000.

The 300-year-old text, which is in Latin, was bought by an unidentified buyer from Europe.

The book was one of 3,000 items auctioned that had belonged to Haskell Norman, a psychiatrist from San Francisco who died two years ago.



US scientists rally behind funding bill

Sir — Read the *Nature* editorial, “Think of a number and double it”, and you will conclude that scientists have lost their savvy — or at least their ability to crunch numbers (*Nature* **393**, 1; 1998). The object of your criticism was the National Research Investment Act of 1998, which has drawn 17 bipartisan co-sponsors in the US Senate and represents many of the sentiments of the unified statement on research supported by the leaders of 110 science, engineering and mathematics organizations. Recognizing the extraordinary importance of science and engineering for the future of the nation, the bill would authorize doubling the federal investment in civilian research over a ten-year period.

As supporters of the legislation, we take strong issue with the facts presented in the editorial and with its logic. The bill, known as S1305, was submitted last fall by Senators Phil Gramm (Republican, Texas), Joseph Lieberman (Democrat, Connecticut), Peter Domenici (Republican, New Mexico) and Jeff Bingaman (Democrat, New Mexico). It aims to accomplish three goals: to help stanch the research budget haemorrhaging of recent years; to raise the visibility of science and engineering within the Senate; and to put the White House on notice that Congress, too, has a strong and enduring interest in the science and technology issue.

The bill has already achieved a measure of these three goals. Last February, the president delivered a strong budget

recommendation on science for fiscal year 1999. In March, the Senate wrote a budget resolution that identifies science as a priority and recently followed that with appropriations allocations that provide opportunities for science account increases. Finally, 17 co-sponsors have joined the bill, and many other senators have become more aware of the nexus between federal research investments and our national prosperity.

The bill has also served as a rallying point for the entire science and engineering community. It has drawn the disparate parts together. And it has made it possible for scientists and engineers around the country to make their case to the public. To date, at least five major newspapers have carried editorials that emphasize the importance of research to the nation's future and to the economic vitality of individual states and localities, all based on the premise of S1305.

To assert, as your editorial does, that S1305 simply carries the message ‘spend’ is to miscast the role of the bill. The legislation follows a time-honoured tradition of sending a message to congressional appropriators that their colleagues want an issue to be taken seriously — in this case research. You seem to have misunderstood how the process works.

Nature's criticism of S1305 also suffers from a serious logical flaw. You assert that the bill's \$170 billion ten-year price tag has not been justified. But you also note that the case has been made convincingly for the

National Science Foundation and the National Institutes of Health (NIH), two agencies that Congress and the White House strongly support. Yet it is the activities of these two agencies that account for most of the growth.

The strength of American science and engineering lies in the multiplicity of agency support. At a time when science and engineering have become so thoroughly interdependent, maintaining the health of this structure takes on even greater significance. Representative John Porter (Republican, Illinois), chairman of the House appropriations subcommittee that oversees the NIH, stated at a hearing on 20 May: “Just as we don't want to set disease against disease, neither do we want to set one science against another.... All have to be valued and brought along at a relatively equal pace.” Those are the goals of S1305, with which we concur. We are mystified that you seemed not to share this view.

Martha Sloan (Chair)

*American Association of Engineering Societies,
1111 19th Street, NW, Washington, DC 20036, USA*

Arthur Jaffe (President)

*American Mathematical Society,
1527 18th Street, NW, Washington, DC 20036, USA*

Paul Walter (President)

*American Chemical Society,
1155 16th Street, NW, Washington, DC 20036, USA*

Andrew M. Sessler (President)

*American Physical Society,
One Physics Ellipse, College Park, MD 20740, USA*

Combinatorial chemistry in the hunt for medicines

Sir — Your feature on bioprospecting (*Nature* **392**, 535; 1998) conveys a misleading impression of combinatorial libraries: “At present, all these synthetic molecules are polymers”. Although early libraries were mainly based on peptides and nucleotides, the current focus¹ is on ‘small molecules’ that are not polymeric and have drug-like physicochemical properties.

I have previously pointed out a fundamental difference between combinatorial chemistry and natural product screening: “At the present time, libraries are constructed by short sequences of synthetic transformations, whereas natural products are the results of much longer and more creative pathways”². Libraries would approach the structural intricacy of natural products if combinatorial syntheses of ten or more steps were practical. Given the current pace of development, this may well be possible in the

future. It is not just a question of increased efficiency, though — chemists string together familiar reactions, while nature is adept at assembling unusual ring skeletons and functional groups.

For example, the chemistry of beta-lactams and enediynes was virtually unknown until their presence was recognized in antibiotics. However, complexity is not always necessary nor desirable. Many medicines are made in a few steps by straightforward reactions, and there is no reason why combinatorial chemistry in its present state cannot lead to many more such drugs.

Anyone engaged in natural product screening will be painfully aware of its bottlenecks. Improvements occur in small increments, unlike the exponential progression of combinatorial chemistry which has led to its enthusiastic reception by the pharmaceutical industry.

The major effect of such comparisons has been to provide a long overdue push to innovate and revolutionize the process of natural product screening, which will

ultimately benefit drug discovery as a whole.

A. Ganesan

*Institute of Molecular and Cell Biology,
30 Medical Drive, Singapore 117609*

1. Balkenhohl, F., von dem Bussche-Huenefeld, C., Lansky, A. & Zechel, C. *Agnew. Chem. Int. Ed. Engl.* **35**, 2288 (1996).
2. Ganesan, A. *Agnew. Chem. Int. Ed. Engl.* **35**, 611 (1996).

Call for change in Brazil

Sir — Fernando C. Reinach addresses important questions on science and technology development in Latin American countries (*Nature* **392**, 647–648; 1998).

A big drawback in our public universities in Brazil is that salaries are not based on productivity. Whether one likes it or not, that is unrealistic in present times. Unfortunately, corporatism is so strong and entrenched in Brazil's academic community that change will only happen if imposed by the government. This, of course, will provoke strikes, and waste time and money.

The problem is that, the more the

academic community resists change, the greater is the risk of losing support from society, and therefore of the flowering of private, profit-making universities. This will have severe consequences for science.

Eurico C. de Oliveira

*Instituto de Biociencias, Universidade de São Paulo,
C postal 11461, 05422-970 São Paulo, SP, Brasil*

Scientists at the sharp end in a disaster zone

Sir — In a News item (*Nature* **392**, 743–744; 1998), you discuss the role of scientists in the management of the Montserrat volcano crisis, on the basis of the verbal presentation of an unpublished study at a meeting.

This study appears to rely mainly on the findings of a student's research project that covered a few months of this prolonged emergency (now nearing three years duration). As far as the Montserrat Volcano Observatory is concerned, the study and visit had no official standing at the time, and was conducted outside the ambit of any of the emergency organizations operating in Montserrat. This work is, as yet, unrevealed for scrutiny and it remains to be seen whether the methods used were robust and reliable, particularly in respect of canvassing accurate opinions from a complex West Indian society under stress. This perspective has been given prominence and currency by your report and, by publishing unsubstantiated assertions about the conduct of the crisis, there is a danger of misleading your readers. On behalf of the scientists who have worked at the observatory throughout the crisis, we reject the main thrust of the piece, and many of its inferences.

Strong emphasis is placed in the report on "predictions" during the eruption, and the ways in which they are purported to have failed. Throughout the crisis, the observatory's scientists have been providing descriptive and probabilistic scenarios of potential eruptive activity over various timescales, using layman's terms as far as possible. They have not, at any time, offered public predictions in any accepted scientific sense of the term. The vast majority of the population developed a sound understanding of the eruption and its implications, and quickly accepted that many aspects of volcanic phenomena are unpredictable.

The examples of supposed predictions cited in the article are either inappropriate or incorrect. For instance, the authorities and the public had been repeatedly advised that explosions, such as that of 17 September 1996, could occur with little or no warning. The heightened alert that accompanied the strong eruptive activity on

19 December 1996 was not a failed prediction. What happened was that observed activity, such as seismicity build-up and ongoing dome collapse, exceeded thresholds for raising the alert level, agreed previously with the civil authorities to indicate a state requiring precautionary evacuation of selected areas. To term this a "false alarm" is tendentious — the warning signs, and official concern, were real and had to be acted upon.

The scientists of the observatory team have provided timely hazard appraisals and detailed risk mitigation strategies, using modern methods of structured assessment, and their work has recently been scrutinized by a House of Commons select committee and by a group chaired by the UK government's chief scientific adviser. In addition to their arduous and often dangerous scientific investigation of the eruption, the team has sustained a public outreach effort probably unsurpassed in any previous volcanic emergency. There is a wealth of documentary evidence about the public's understanding and response to the scientific component of this emergency, which will become increasingly available once the crisis has subsided, and upon which a balanced and complete account of the observatory's contribution to the management of the Montserrat eruption can be based. Only then can successes or shortcomings be rationally judged.

In particular, all the factors surrounding the 19 fatalities on 25 June 1997 (and two others later on) are yet to be fully elucidated and appraised. We wish to stress that scientific warnings had been issued (or acted upon) between December 1996 and June 1997. Although some people had been allowed back into certain areas after the December activity subsided, a re-evacuation and extended exclusion zone were implemented well before the big dome collapse occurred. All the casualties occurred within that exclusion zone, in areas that had been clearly identified as high risk. This casualty toll should be placed in its wider context: at the outset of the crisis in July 1995, when many in Montserrat were unaware that the Soufriere Hills was an active volcano, there were about 7,000 people living and working in areas that are now buried by pyroclastic flows or obliterated by violent surges. Their lives were successfully protected, albeit at great personal cost and privation.

In an editorial (*Nature* **388**, 1; 1997), you used the tragic events of 25 June 1997 to enjoin scientists confronting natural disasters to co-operate more closely with social scientists. In principle, this is a sensible proposal, and worthwhile if there is scope for achieving real benefit. But in a free society there will always be individuals who will not heed advice or instructions to avoid

hazardous situations such as those on the flanks of an erupting volcano. This was a major factor in the Montserrat fatalities, not the minutiae of communications between scientists and the public. Many people put themselves in danger attempting to convince the victims to leave those areas and, in terms of any scientific function, it is not clear what more could have been done to avert these casualties.

The fact that a significant number of individuals may knowingly choose to accept extreme levels of risk to remain in a danger zone will present a major challenge for societal risk management at volcanoes elsewhere. A careful, collaborative analysis of all the experience from the Montserrat crisis is merited, not a unilateral and ill-considered rush to judgement. Your article and its shortcomings are, therefore, at best unhelpful, and may actually discourage hard-pressed scientists (endeavouring to help protect life and limb) from engaging with colleagues from the social sciences in any similar safety-critical situation.

Willy Aspinall, Peter Francis, Lloyd Lynch, Richard Robertson, Keith Rowley, Steve Sparks, Simon Young and others

*Montserrat Volcano Observatory,
St Johns, Montserrat, West Indies
e-mail: mvo@candw.ag*

Sir — A recent News story about the earthquake on Montserrat referred to a report with which I was involved (*Nature* **392**, 743–744; 1998). The thrust of our findings was that the scientists involved have been in the unenviable position of *de facto* disaster managers, a role for which they were not trained nor, it could be argued, should have been involved in.

There has therefore been an additional burden to their already complex and difficult task of being both authors and messengers of volcanic forecasting and warning. It should be recognized that, if the scientists had not been involved, the death toll on the island would have certainly been higher. Their role has been one of crisis management in extreme circumstances.

It was not part of our findings that many residents felt "the scientists' prediction record was sometimes no better than their own". Although one or two may have held this view, this generalization was not made. The credibility of the scientists has generally been high throughout the crisis; indeed, the constant focus on the scientists' activities would appear to be one of the major contributing pressures on them. The project was funded by the UK Department for International Development as part of a two-year project looking at forecasting and warnings in four parts of the world.

David Sanderson

*154 Banbury Road, Kidlington,
Oxford OX5 2BY, UK*

When is a bird not a bird?

Kevin Padian

Birds were once thought to have a large number of features exclusive to the group. One by one those features have also been identified in fossils of certain theropod dinosaurs. Now feathers join the list.

Among all living creatures, only birds have feathers. So when the fossil of a single isolated feather was found in the Late Jurassic rocks of Bavaria shortly after Darwin published *On the Origin of Species* in 1859, it was enough to demonstrate that, astonishingly, birds must have existed in those remote times (now known to be some 150 million years ago). When the skeleton of *Archaeopteryx* was discovered in the same area in 1861, it fulfilled that plumose prediction. But the skeleton was recognized as a bird's mainly because it had feathers on its wings and tail. Nothing else except an odd, boomerang-shaped wishbone seemed to ally it to living birds. In fact, some other specimens were assigned to *Archaeopteryx* only belatedly, because no feathers were preserved in the fossils concerned. Those specimens had instead been taken for small carnivorous (theropod) dinosaurs.

On page 753 of this issue, Ji *et al.*¹ show unequivocally that clothes don't make the

bird. They describe two small theropod dinosaurs from geological beds in the Liaoning province of China¹⁻⁴. The age of the beds is disputed; although initial reports suggested they are Late Jurassic, radiometric dates and other evidence now point to the Early Cretaceous⁴ — that is, around 145 million years or maybe later. These theropods have both plumulaceous (down-like) and vaned, barbed feathers on the body, arms, legs and tail. But these animals were clearly not birds, and they were clearly not capable of flight.

One form, *Protarchaeopteryx*, has been briefly described in a Chinese journal². It has down-like feathers on its body and tail, and vaned, barbed, symmetrical feathers along at least the end of the tail, in a fan-like pattern. The second, *Caudipteryx*, has remiges (primary feathers) attached to the second (longest) finger of the hand, though the arms are much shorter than in birds. These feathers are also vaned and barbed, and down-like feathers are also preserved. Were preserva-

tion more complete, we would have an even fuller idea of the plumage of these animals. This is the most that can be said at the moment, but new specimens continue to emerge from this locality.

If these are true feathers — and it can scarcely be doubted — what does this do to our conception of birds? The bottom line is that it simply forces us to revise our idea of the association of feathers with the animals we call birds. This is why.

Systematists *define* the names of organisms (taxa) by their ancestry; in this case, birds (Aves) consist of *Archaeopteryx* plus living birds (Neornithes) and all the descendants of their most recent common ancestor (Fig. 1). But we *diagnose* (recognize) birds by unique features that only they possess; these are inherited from that most recent common ancestor. (Ji *et al.*¹ give three such features, none of which relates to feathers.) So the diagnostic characteristics follow from our definition of ancestry. Why worry if feathers turn out to be shared by a wider group than birds alone? We still define birds as *Archaeopteryx* and its later relatives. *Protarchaeopteryx* and *Caudipteryx* may have feathers but they're not birds, because they're not members of that ancestral club of *Archaeopteryx* and living birds. Ji *et al.*¹ show that these animals belong to a group of dinosaurs known as the maniraptoran coelurosaurs, which include the small theropods most closely related to birds. However, their analysis of evolutionary relationships does not encompass a broader group of coelurosaurs, which might eventually help to pin down the position of *Protarchaeopteryx*.

By admitting that plumage did not first spring full-blown on the wings of *Archaeopteryx*, we are free to examine how feathers evolved in the first place. At the 1996 meeting of the Society of Vertebrate Paleontology, Dr Chen Pei-ji disclosed photographs of an amazing specimen — a small theropod dinosaur, also from the Liaoning beds, with a fringe of hair-like or down-like structures along its neck and backbone. This animal, dubbed *Sinosauropteryx* by Ji and Ji^{3,4}, turns out to be closely related to the basal (relatively 'primitive') coelurosaur *Compsognathus*, a chicken-sized, short-armed form found in the same deposits as *Archaeopteryx*⁵. Without even seeing the specimens, critics of the theropod–bird connection tried to pass off these remains as artefacts of preservation. They claimed that the structures were frayed internal collagen fibres, not external or epidermal in origin at all, and that similar structures in sea snakes suggested that these theropods lived in an aquatic environment⁶. Those doubts can now be put to rest^{3,4}.

The evolution of carnivorous dinosaurs through basal coelurosaurs into birds shows some unmistakable trends⁷. As their evolutionary history has become better known, we have seen wishbones, breastbones, hollow

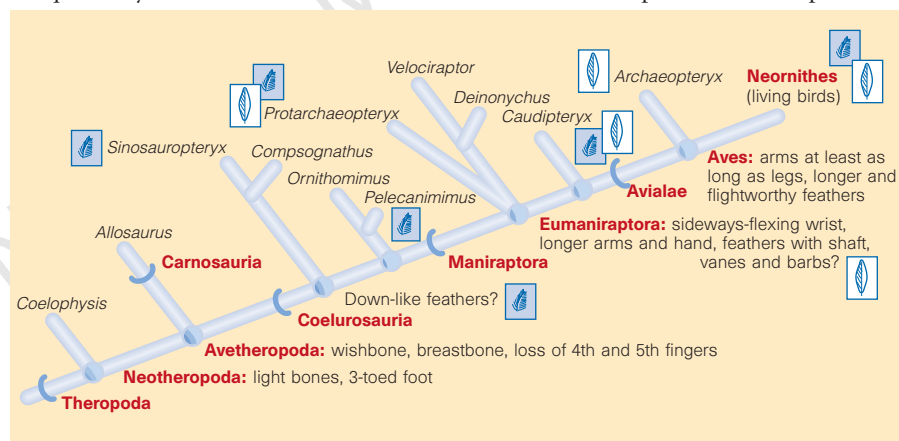


Figure 1 Distribution of known feather types and other features in selected groups of theropod dinosaurs, including birds (Aves). The nodes (junctions) on the diagram provide the names and shared features of the groups connected above the nodes. Hence, the Avetheropoda (everything but *Coelophysis* in this diagram) have wishbones and breastbones and have lost the fourth and fifth fingers. The fossils of *Protarchaeopteryx* and *Caudipteryx* have both the vaned and down-like feathers seen in living birds, so we infer that their common ancestor (group Eumaniraptora) had them. (This may be true for the *Velociraptor*–*Deinonychus* group, by inference, unless *Protarchaeopteryx* later turns out to be more closely related to birds than they are.) Down-like feathers are not known in specimens of *Archaeopteryx*. But the base of its vaned feathers are down-like⁹, like those of *Protarchaeopteryx* and living birds, and the reasonable inference is that *Archaeopteryx* did in fact possess them. *Sinosauropteryx* has only the beginnings of down-like feathers, which could be seen as an aberration or a foreshadowing of their later elaboration in the Eumaniraptora, but their similarity to down-like feathers is very strong⁴. Similar structures have been reported in the ornithomimid *Pelecanimimus*¹⁰, but these now appear to be other soft-part structures¹¹.

bones, long arms and hands, sideways-flexing wrists, nesting behaviour and rapid growth rates all disappear from the avian catalogue of exclusive features. Like feathers, as Ji and colleagues' work¹ shows, these features first evolved in coelurosaurs for purposes completely unrelated to flight⁷. The proto-feathery fringe of *Sinosauropteryx*, now known to extend to its flanks as well as along its midline⁴, was obviously not made for flight and it is questionable whether it could have served any kind of function in insulation. Camouflage, display and species recognition come to mind as other possibilities⁸.

Caudipteryx and *Protarchaeopteryx* go it one better, evolving long feathers with a central rachis (shaft). Were these feathers airworthy? Their vanes are symmetrical and very even, suggesting interlocking barbs, although most flying birds have asymmetrical feathers. However, the arms of *Caudipteryx* are only 60% as long as the legs, and they are only two-thirds as long in *Protarchaeopteryx* (in *Archaeopteryx* they are of more or less equal length). Evidently, the arms and feathers were not large enough for flight (the plumage is not yet known well enough to say if it was effective as insulation, or what other functions it may have served). The feathers of *Archaeopteryx* seem to be a natural extension of this trend, so to speak,

but they are not much different qualitatively. So the available evidence suggests that structurally airworthy feathers may have evolved before they were long enough, or their possessors able to use them, for flight.

The work of Ji *et al.*¹ should lay to rest any remaining doubts that birds evolved from small coelurosaurian dinosaurs. These new discoveries will excite the public and scientists alike by showing that down-like and later vaned body feathers evolved before flight feathers, and that a full complement of feathers was present in coelurosaurs before birds were invented. □

Kevin Padian is in the Department of Integrative Biology and the Museum of Paleontology, University of California at Berkeley, Berkeley, California 94720-3140, USA.

e-mail: kpadian@socrates.berkeley.edu

1. Ji, Q., Currie, P. J., Norell, M. A. & Ji, S.-A. *Nature* **393**, 753–761 (1998).
2. Ji, Q. & Ji, S.-A. *Chinese Geol.* **238**, 38–41 (1997).
3. Ji, Q. & Ji, S.-A. *Chinese Geol.* **233**, 30–33 (1996).
4. Chen, P.-j., Dong, Z.-m. & Zhen, S.-n. *Nature* **391**, 147–152 (1998).
5. Ostrom, J. H. *Zitteliana* **4**, 73–118 (1978).
6. Geist, N. R., Jones, T. D. & Ruben, J. A. *J. Vert. Paleont.* **17**, 48A (1997).
7. Padian, K. & Chiappe, L. M. *Biol. Rev.* **73**, 1–42 (1998).
8. Cowen, R. & Lipps, J. H. *Proc. 3rd N. Am. Paleont. Conv.* (Montreal) 109–112 (1982).
9. Griffiths, P. J. *Archaeopteryx* **14**, 1–26 (1994).
10. Pérez-Moreno, B. P. *et al.* *Nature* **370**, 363–367 (1994).
11. Briggs, D. E. G. *et al.* *J. Geol. Soc. Lond.* **154**, 587–588 (1997).

Fullerenes

C₆₀'s smallest cousin

James R. Heath

In 1990, Kratschmer, Lamb, Fostiropoulos and Huffman (KLFH)¹ caught the chemical world off-guard when they announced their synthesis of C₆₀-buckminsterfullerene. It was surprising not only that bulk quantities of fullerenes could be prepared, but also that the process was so simple. In chemical terms, the KLFH synthesis is essentially a single-step procedure — strike an arc across two graphite electrodes and large quantities of fullerenes are produced.

In reality, however, the formation of fullerenes proceeds through an extremely fast and complex sequence of reactions, starting with carbon atoms and dimers². In spite of the tremendous advances in fullerene chemistry since the initial discovery of C₆₀ in 1985 (ref. 3), very little detailed insight into the formation mechanisms of the fullerenes and related structures has been achieved. On page 771 of this issue⁴, Piskoti, Yarger and Zettl present evidence for the formation of bulk quantities of C₃₆-fullerene. This represents the first direct evidence that a fullerene smaller than C₆₀ is produced in the KLFH synthesis. The very fact that C₃₆ is formed implies much about the chemistry that leads to C₆₀ and the higher fullerenes.

Most researchers agree that the arc-gen-

erated vapour of atoms and dimers initially condenses to form linear carbon chains containing from 3 to 20 atoms, and the importance of these structures was realized nearly 40 years ago⁵. The next stage in fullerene growth involves a transition, somewhere between C₁₀ and C₂₀, of carbon chains into

monocyclic ring structures. Monocyclic rings have also been the subject of attention for many years, and ion-mobility measurements indicate that ring-like structural isomers may actually exist up to C₆₀ or higher⁶. It is also clear that more compact structures exist for clusters larger than C₃₀, and those structures are probably the precursors to C₆₀ and the higher fullerenes.

Two growth mechanisms have been proposed to explain the KLFH fullerene synthesis, and they predict very different types of structures for the C₃₀ to C₅₉ size range. The first mechanism, the pentagon road, postulates that three-dimensional structures containing hexagons and pentagons become important above C₃₀ (ref. 7). The rule of thumb for these clusters is that the number of adjacent pentagons is always minimized, and so these structures tend to remain open as they grow. The smallest closed fullerene is the first one that can avoid adjacent pentagons, and is, of course, C₆₀.

In the second proposed mechanism, the fullerene road, three-dimensional structures containing hexagons and pentagons also become important somewhere around C₃₀ (ref. 8). Adjacent pentagons are allowed along this road — it is dangling bonds that are minimized.

If a cluster has no adjacent pentagons (as in C₆₀), then it is kinetically stabilized. For the pentagon road, dangling bonds at the edge of the carbon clusters provide reactive sites for cluster growth. For the fullerene road, adjacent pentagons, which are the points of highest strain energy for a closed fullerene, provide reactive sites for growth through addition of C₂ units. Both proposed mechanisms lead naturally to large amounts of C₆₀, C₇₀ and the higher fullerenes. But whereas thermodynamic arguments favour the fullerene road⁹, kinetic arguments favour the pentagon road.

Manolopoulos and Fowler¹⁰ have predict-

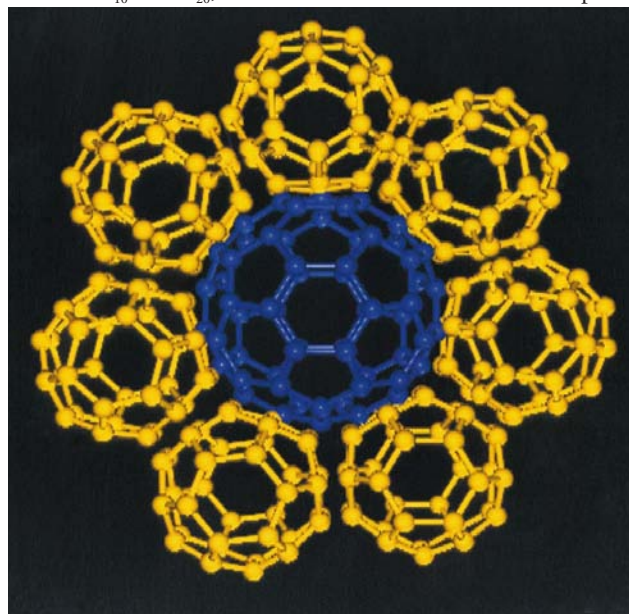


Figure 1 Little and large relations — C₆₀ surrounded by seven of its smaller C₃₆-fullerene cousins, which Piskoti *et al.*⁴ have prepared in bulk form. C₃₆ is golden yellow in solution, whereas C₆₀ is blue/purple. The apparent symmetry of C₃₆ implies much about C₆₀'s chemical ancestry. In particular, C₃₆ has the minimum possible number of shared pentagon–pentagon bonds, which is consistent with the growth mechanism for C₆₀ known as the fullerene road. (Figure courtesy A. Zettl.)

ed a lowest-energy fullerene road, starting at C_{24} (which has only one structural isomer) and terminating at C_{60} . At each step, the closed (fullerene) isomer that was characterized by the smallest number of adjacent pentagons was generated either directly by C_2 addition to the next smallest fullerene, or through a Stone–Wales rearrangement¹¹ which provides a mechanism for pentagons and hexagons to switch places. Although most fullerenes smaller than C_{60} are characterized by many possible structural isomers, the lowest-energy pathway on the fullerene road predicts that there are only a small subset of structures for all fullerenes smaller than C_{60} .

Until now, this issue has been largely academic. The difference between these two mechanisms involves a series of (presumably) highly unstable carbon clusters that have eluded structural characterization. However, the announcement by Piskoti, Yarger and Zettl⁴ that bulk amounts of C_{36} can be prepared using the KLFH synthesis provides strong evidence that the thermodynamically favoured pathway, the fullerene road, dominates bulk fullerene production. Measurements of solid-state C_{36} by nuclear magnetic resonance indicate that it probably has D_{6h} symmetry which, among other things, implies that it has a six-fold rotation axis. This result is also consistent with the fullerene road. The D_{6h} form of C_{36} is characterized by the minimum possible number of shared pentagon–pentagon bonds (12) (Fig. 1), and is one of two low-energy structures predicted for C_{36} -fullerene. Existing electronic structure predictions for the D_{6h} structure indicate that it should be characterized by a triplet ground electronic state. It is a little surprising that this structure is made in bulk form, rather than the energetically competitive and lower-symmetry D_{2d} structure, which has a singlet ground state¹². However, the predictions are based on low-level Hückel theory, which doesn't account for interactions between various molecular orbitals. A more sophisticated approach that does account for such interactions (such as the mixing of sigma and pi bonds) may well predict a singlet D_{6h} ground state. In addition, C_{36} apparently reacts quickly in air to produce $C_{36}H_6$, and that structure is almost certainly a singlet molecule.

Is C_{36} the smallest fullerene that is likely to be made in bulk form? Probably yes. The molecule is surely characterized by a tremendous amount of curvature-induced bond strain, and so it is highly improbable that an even smaller fullerene will be isolated. But it now seems possible that other fullerenes between C_{36} and C_{60} could be prepared. One attractive target is C_{50} — the D_{5h} structure of C_{50} -fullerene is the most stable C_{50} isomer, and it is also a dead-end on the lowest-energy pathway on the fullerene road¹⁰. □

James R. Heath is in the Department of Chemistry and Biochemistry, University of California at Los

Angeles, 405 Hilgard Avenue, Los Angeles, California 90095-1569, USA.
e-mail: Heath@chem.ucla.edu

1. Kratschmer, W., Lamb, L. D., Fostiropoulos, K. & Huffman, D. R. *Nature* **347**, 354–356 (1990).
2. Meijer, G. & Bethune, D. S. *J. Chem. Phys.* **93**, 7800–7802 (1990).
3. Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. & Smalley, R. E. *Nature* **318**, 162–164 (1985).
4. Piskoti, C., Yarger, J. & Zettl, A. *Nature* **393**, 771–774 (1998).
5. Pitzer, K. S. & Clementi, E. *J. Am. Chem. Soc.* **81**, 4477–4490 (1959).
6. Hunter, J., Fye, J. & Jarrold, M. F. *J. Phys. Chem.* **97**, 3469–3471 (1993).

7. Smalley, R. E. *Acc. Chem. Res.* **25**, 98–105 (1992).
8. Heath, J. R. in *Fullerenes* ACS Symp. Ser. 481 (eds Hammond, G. S. & Kuck, V. J.) 1–21 (Am. Chem. Soc., Washington DC, 1992).
9. Bates, K. R. & Scuseria, G. E. *J. Phys. Chem. A* **101**, 3038–3041 (1997).
10. Manolopoulos, D. E. & Fowler, P. W. in *The Chemical Physics of the Fullerenes: 10 (and 5) Years Later* (ed. Andreoni, W.) 51–70 (Kluwer, Dordrecht, 1996).
11. Stone, A. J. & Wales, D. J. *Chem. Phys. Lett.* **128**, 501–504 (1986).
12. Fowler, P. W. & Manolopoulos, D. E. *An Atlas of Fullerenes* (Clarendon, Oxford, 1995).

Bioluminescence

Dolphins glow with the flow

Peter J. Herring

The dynamics of flow over the surface of animals — ranging from insects in flight to swimming fish — have been examined using optical methods and markers such as dye, air bubbles and smoke. At the same time, amateur and professional seafarers (and biologists) have considered themselves extremely lucky if they have seen the firework displays produced by dolphins swimming in waters that contain high levels of bioluminescent plankton (Fig. 1). Now, physicist Jim Rohr, biologist Michael Latz and their colleagues have ingeniously combined the two approaches, and they report their findings in the *Journal of Experimental Biology*¹. The authors analysed the light output that is induced from natural populations of bioluminescent dinoflagellates by free-swimming dolphins, and they have used the results to interpret some of the characteristics of flow over these animals.

Bioluminescence is the remarkable ability of many living organisms to produce visible light. These organisms, which include

bacteria and dinoflagellates, fireflies and fish, use special chemiluminescent reactions to produce energy in the form of light. The luminous dinoflagellates are unique flow markers, lighting up only when particular levels of shear stress are present.

Previous attempts to use anecdotal accounts and sketches of dolphin-stimulated bioluminescence to infer the flow field around an animal's body have foundered in the absence of hard data. These attempts have led, for example, to the incorrect assumptions that the presence of the bioluminescence indicates turbulent flow and its absence from particular regions implies laminar flow.

Rohr *et al.*¹ have now meticulously integrated information from field studies, involving image-intensified videos of the bioluminescence stimulated by two trained dolphins, with computational flow models and laboratory experiments that quantify the sensitivity of the dinoflagellates to flow stimulation. The results indicate that

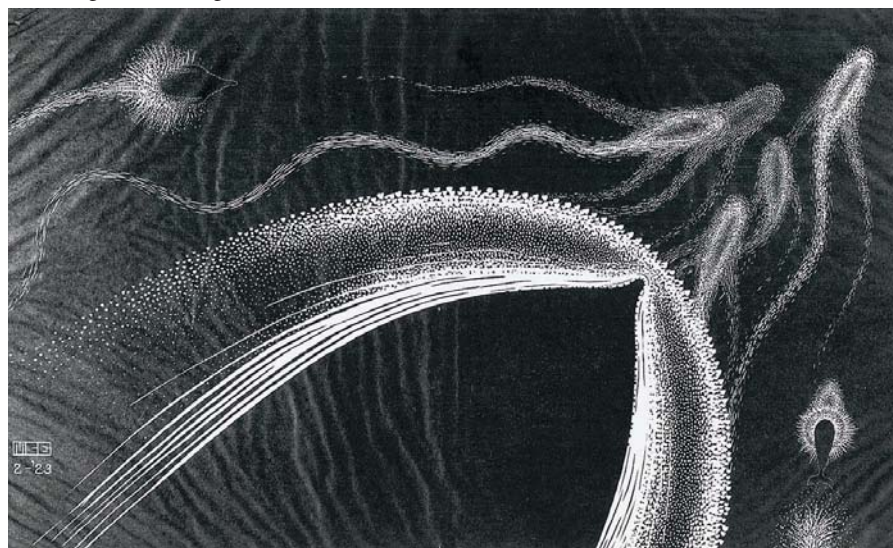


Figure 1 'Dolphins in phosphorescent sea'. The inspiration for this woodcut, created by M. C. Escher in 1923, was the flow-induced bioluminescence that occurs on dolphins when they swim through waters that contain high levels of bioluminescent plankton. Rohr and colleagues¹ used the phenomenon to analyse the flow around these animals.

CORDON ART

dinoflagellate bioluminescence is initiated by *both* laminar and turbulent flows, whenever the levels of shear stress are greater than 0.1 N m^{-2} . Using a model ellipsoid in flow conditions similar to those for a dolphin of equivalent size gliding at 2 m s^{-1} , the authors calculated that the contour of this level of shear approximately demarcates the edge of the boundary layer. This is the layer where the mean flow velocity is reduced to less than 99.5% of that in the free stream beyond. Thus, the dolphin will induce bioluminescence within the boundary layer whether the flow is laminar or turbulent.

Next, Rohr and his colleagues temporarily attached a Velcro band or a cup to their dolphins. The video data showed changes in the bioluminescence pattern, and they attributed this to flow separation or forced transition from laminar to turbulent flow. They also found consistently low levels of bioluminescence over the melon (the rounded bump containing waxy material at the front of the dolphin's head) and on the leading edges of the fins, where the surface shear stress is predicted to be highest. They conclude that these variations in the intensity of bioluminescence do not indicate differences in flow type but, rather, differences in the thickness of the boundary layer. This determines the volume of flow in which the dinoflagellates experience above-threshold levels of shear stress.

The dynamics of flow around bumblebees, birds and trout can be followed under experimental conditions in the laboratory. It is much more difficult to conduct meaningful experiments on large, fast-moving, aquatic animals, but Rohr and colleagues' study shows one promising method of flow visualization. Potential candidates for this technique include billfish, seals, penguins and even the smaller whales. The effect is produced by almost any moving object and, in the past, the bioluminescence produced by torpedoes, submarines and surface ships has focused considerable attention on the phenomenon. Indeed, it has been one approach to the non-acoustic detection of submarines², and the last U-boat sunk in the First World War was located by its bioluminescent trail.

From the dolphin experiments, a vehicle design that minimizes the thickness of the boundary layer would seem to present the least visible signal at a given speed to an observer or sensor. The (usually unacceptable) alternative is simply to reduce the shear stress by going more slowly. However, the immediate boundary layer over the vehicle is the source of only a small proportion of the total light. Most of the light is generated by flow separation of the boundary layer into the much larger turbulent volume of the wake. One patent application even sought to reduce the luminescent signal from a ship by pumping detergent into the water, in the

hope of decreasing the turbulent flow along the hull! The dolphin's design and method of propulsion leaves a very different turbulent wake than does a torpedo or submarine; nevertheless, in wartime the luminous contrails of dolphins have often given seafarers cause for concern rather than admiration.

A dolphin's luminous trail probably does not seriously increase its risk of predation. But the same cannot be said for other animals — in nature, the bioluminescence of dinoflagellates seems to act as a 'burglar alarm' against zooplankton predators, giving them (and other animals) away to their own predators³. Similar luminous signals, identified by spotter aircraft, have been used to target schools of fish and krill near the surface, leading to the suggestion^{4,5} that this could be a means of surveying the populations of krill in the Southern Ocean. The dolphin experi-

ments reported by Rohr *et al.*¹ provide an excellent example of the value of these luminous dinoflagellate markers, and they have other promising applications on both large and small scales. For oceanographers, they can illuminate features such as fronts, solitons (solitary waves) and wind shear, whereas for bio-engineers they can show the flow in systems such as biological reactor vessels and artificial heart chambers. □

Peter J. Herring is at the Southampton Oceanography Centre, Empress Dock, Southampton SO14 3ZH, UK.

e-mail: p.j.herring@soc.soton.ac.uk

1. Rohr, J., Latz, M. L., Fallon, S., Nauen, J. C. & Hendricks, E. *J. Exp. Biol.* **201**, 1447–1460 (1998).
2. Stefanick, T. *Sci. Am.* **258**, 25–31 (1988).
3. Fleisher, K. J. & Case, J. F. *Biol. Bull.* **189**, 263–271 (1995).
4. Cram, D. L. & Schulein, F. H. *J. Cons. Int. Explor. Mer* **35**, 272–275 (1974).
5. Cram, D. L. & Malan, O. G. S. *Afr. J. Antarct. Res.* **7**, 14–19 (1977).

Oceanography

Pumping iron makes thinner diatoms

Ed Boyle

In 1988, the late John Martin suggested that phytoplankton productivity in 'high nitrate, low chlorophyll' (HNLC) provinces was limited by the iron supply¹. Work on this topic has since come a long way, and the latest instalments appear in work published two weeks ago by Hutchins and Bruland², and by Takeda on page 774 of this issue³. Silica now enters the equation.

Although iron's importance for marine nutrition had been understood well before the late 1980s (see, for instance, ref. 4), Martin's group was the first to test the 'iron hypothesis' with appropriately clean and sensitive methods in sea water. In subsequent years, numerous studies have shown that nanomolar additions of iron to natural phytoplankton communities increases productivity and photosynthetic electron-transport efficiency. This result is seen both in bottle incubation experiments on board ships and in mesoscale experiments, on the kilometre level, *in situ*⁵. Iron limitation has been shown to operate in the subarctic Pacific, the high-latitude Southern Ocean, and the eastern equatorial Pacific. It has also been suggested that iron availability may limit nitrogen fixation in nutrient-poor ocean regions⁶.

Hutchins and Bruland² and Takeda³ show that iron affects diatom uptake of silica relative to nitrate (in one of the studies³, relative to phosphorus also; and by inference in both studies, relative to carbon as well). The results suggest that iron has a much broader role in regulating oceanic biogeochemical cycles than has been appreciated. These cycles may have operated quite differently during 'dusty' ice-age periods, when there

was likely to have been much greater wind-borne input of iron from land to the surface ocean than there is now.

Hutchins and Bruland² examined the effects of iron addition to a coastal upwelling environment (the California Current, northern California, Big Sur region). The initial goal of their study was to find out whether productivity in this province is limited by iron availability. Normally, coastal regions are not expected to be iron-limited because rivers and continental-shelf sediments provide an ample supply of the element. However, in the Big Sur region, low fluvial input and a narrow continental shelf minimize these sources of iron. The intense supply of upwelled nutrients and low iron input results in HNLC conditions close to the California coast.

Hutchins and Bruland's experiments showed that addition of iron increases the reproduction rates and nitrate uptake of natural phytoplankton communities incubated for several days in the laboratory. Adding iron to samples taken from nearby sites with naturally higher levels of iron did not further increase growth. This result establishes the importance of iron in a fourth oceanic province.

However, there is a new twist to the story. In all of the previous experiments, as well as in this new work, increased productivity is almost entirely due to the response of large diatoms — one characteristic of which is their siliceous skeleton. Previously, however, the effect of iron enhancement on silica uptake was not monitored. In these new experiments it was, and the result is

surprising. Although the diatoms were reproducing and taking up nitrate more quickly, the effect of this extra growth on silica uptake was minimal. In consequence, Si:N (and by inference, presumably also Si:C and Si:P) ratios were two to three times as high in the unfertilized controls as in the iron-enhanced cultures.

Takeda³ has extended silica monitoring to the three previously established HNLC provinces, and has also isolated two species of Antarctic diatom in culture to examine the specific response of these species in addition to that of the complex natural community. In all cases, he found the same result as Hutchins and Bruland — extra iron resulted in enhanced diatom reproduction rates and nitrate uptake, but only slightly increased uptake of silica. He also measured Si:P ratios, and verified that Si:P ratios were lower in the iron-addition experiments.

The two papers^{2,3} show that iron-limited diatoms grow thicker silica shells. When they die, these diatoms can be expected to sink faster, lose less silica on their descent to the sea floor, and so be better preserved on the sea floor in the form of opal deposits (which consist of hydrated SiO₂). Opal accumulation rates are used to trace past levels of productivity, and the new results suggest that the approach may be biased. This is because a dusty glacial atmosphere may have supplied more iron to the oceans and increased diatom carbon export to the sea floor without greatly enhancing silica export. If this was so, higher productivity would not be reflected by higher accumulation rates of sedimentary opal and would eliminate an apparent contradiction — the lack of observed higher accumulation rates of opal had been a sticking point for the hypothesis that enhanced Southern Ocean productivity was a factor driving reduced levels of atmospheric CO₂ during glacial periods^{7,8}.

A further consequence of these findings is the implication that the distribution of silica, nitrate and phosphate within the ocean may depend on iron. In today's ocean, silica is more rapidly removed into the deeper parts of the ocean than the more labile nutrients nitrate and phosphorus. This situation may not be normal, and may occur only during relatively infrequent, dust-starved, warm periods such as the Holocene (the climate period of the past 10,000 years). Over the Pleistocene, the past two million years, periods of higher dust-flux have been more typical.

So, in the past, diatoms may have been less silicified, resulting in less efficient transport of silica to the bottom of the ocean, with more rapid dissolution and recycling of the silica skeletons that survived transit. The vertical and inter-ocean contrast of silica compared to nitrate and phosphorus may have been less during glacial periods than it has been in warm intervals such as the

past 10,000 years. Furthermore, in glacial periods, higher uptake of nitrogen and phosphorus in areas that are currently HNLC would have affected the distribution of these elements in the surface and upper ocean.

With these new findings, the power of an element — iron — that occurs at low concentrations in the ocean (10⁻⁹ mol kg⁻¹ and less) is becoming more evident. It appears that iron influences the oceanic distributions of nutrients at the level of micromolar concentrations, and also influences the preservation of opaline fossils on the sea floor. Because the distribution of other chemical species (some — such as zinc — with biological importance of their own) are closely cou-

pled to the oceanic cycles of these nutrients, we can expect more surprises as the biogeochemistry of trace metals is investigated further. □

Ed Boyle is in the Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, E34-258, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.

e-mail: eaboyle@MIT.EDU

1. Martin, J. H. & Fitzwater, S. E. *Nature* **331**, 341–343 (1988).
2. Hutchins, D. A. & Bruland, K. W. *Nature* **393**, 561–564 (1998).
3. Takeda, S. *Nature* **393**, 774–777 (1998).
4. Harvey, H. W. *J. Mar. Biol. Assoc. UK* **22**, 205–219 (1937).
5. Coale, K. H. *et al.* *Nature* **383**, 495–501 (1996).
6. Falkowski, P. G. *Nature* **387**, 272–275 (1997).
7. Martin, J. H. *Paleoceanography* **5**, 1–13 (1990).
8. Charles, C. D., Froelich, P. N., Zibello, M., Mortlock, R. A. & Morley, J. J. *Paleoceanography* **6**, 697–728 (1991).

Evolutionary biology

Genitally does it

Darryl T. Gwynne

The genitals of male insects bear a huge number of intricate bits and pieces. The genitalic terms for orthopterans alone (crickets and allies) include phalli, epiprocts, paraprocts, cerci, gonotremes and even titillators (from the Latin *titillo*, 'tickle').

But why have the male's genitals evolved to be so complex? Two answers have been proposed. First, the species-diagnostic nature of these male characteristics suggests that species differences in genitals may have evolved as barriers to insemination, preventing the production of low-quality hybrid offspring. So, in zones of species overlap, the female genital 'locks' of each species diverged, imposing Darwinian selection among conspecific males for a proper fit of each genital 'key'. The second answer is that genital structures diverged through the sexual selection that occurs after insemination for any device that increases fertilization success. Possible examples are claspers, which hold the struggling female until insemination occurs¹, or titillating devices that deliver 'internal courtship signals' to the female. Sperm from males with the best

signals are used in preference to stored sperm from the females' previous mates, and such copulatory discrimination is favoured in females because they pass on the advantageous genital trait to their sons².

On page 784 of this issue, Göran Arnqvist³ provides a rigorous test of the lock-and-key and sexual-selection hypotheses. The rigour is in his methods. First, he measures structural divergence geometrically, thereby overcoming the inevitable subjectivity problems that occur when researchers assess morphological complexity (see, for example, ref. 4). Then he compares these divergence measures for matched pairs of insect groups that share a common ancestor but differ in the pattern of female mating. One group contains species reported to have females that mate only once (monandrous species), whereas the other has multiply-mating females (polyandrous species). If post-insemination selection is important, the genitals should be more divergent in the polyandrous groups because competition between the ejaculates of rivals can occur only when females take

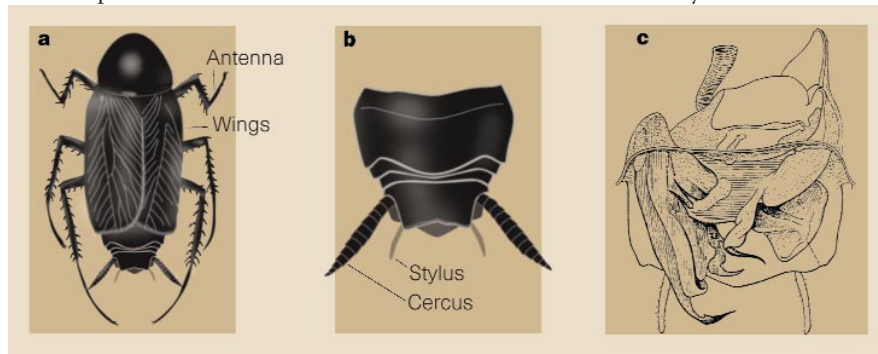


Figure 1 Dorsal views of the 'Swiss army knife' gadgetry in the genitals of a male oriental cockroach, *Blatta orientalis*. Arnqvist³ has studied why such complex male genitals have evolved in insects, and he concludes that it is a result of sexual selection for genital devices that favour a fertilization bias for the mating male's sperm over the stored sperm of the female's previous mates. a, Whole cockroach, from ref. 12; b, c, enlarged genital area from ref. 13.

on additional partners. Conversely, the lock-and-key hypothesis predicts greater genital divergence in monandrous species, because of the potentially large cost to the female if her single copulation is with the wrong species.

Arnqvist's results overwhelmingly supported post-insemination selection. In 18 of 19 pairs sampled from mayflies, flies, beetles and lepidopterans, the shape of male genitals from species in the polyandrous groups has diverged more than in single-mating species. Moreover, this divergence was found only in genitals, and not in other body parts such as legs.

Could any other selective forces explain the differences in genital divergence between the monandrous and polyandrous groups? After all, the two groups do differ with respect to other aspects of their life history — variables that are not controlled for in a comparative study such as Arnqvist's. For example, compared with polyandrous species, monandrous females tend to have a lower fecundity⁵ and they copulate immediately after emerging as adults⁶. Perhaps complex gadgetry is not needed to engage the pliant reproductive parts of a newly emerged female whose integument has not yet hardened. However, hypotheses such as this are unlikely to provide general explanations for the patterns observed by Arnqvist.

One hypothesis that may contribute to these patterns involves a combination of the two models considered by Arnqvist. Hybridization costs do exist in species with polyandrous females, and some post-insemination selection may be a response to these costs. In a beetle, lacewing, fruitfly, grasshopper and cricket (refs 7–9 and refs therein) there is clearly no lock-and-key isolating mechanism — sperm from the wrong species (or, in one case, subspecies) can find its way into the sperm storage organs of females, and even successfully fertilize eggs. But this occurs only when it is the sole ejaculate present, and such heterospecific sperm usually lose the fertilization battle when in competition with sperm from the same species as the female. There is some evidence that this outcome is partly due to the ability of the female's reproductive tract to recognize and favour sperm from her own species. Such cryptic communication of species identity is not well understood⁹.

So here's the alternative hypothesis. Could male genital modifications in some polyandrous species include devices that signal to the female that conspecific sperm is being delivered (similar to the way in which genital structures are thought to signal the quality of a conspecific male²)? This idea is worth investigating, even though it is not a general explanation. For example, it does not explain why divergent genitals exist when there is no threat of mis-mating — such as in isolated island species or in parasites that do

not share hosts with related species². To work out the importance of genital signalling (of any sort) on the evolution of genitalic complexity, we need many more experimental and observational studies on individual species¹⁰.

Finally, it is worth noting that Darwin, who was a very competent entomologist, was keenly aware of “the complex appendages at the apex of the abdomen in male insects”, structures that he thought could function as mate-holding devices¹. He recognized that sexual selection could produce obscure genitalic traits as well as the more widely acclaimed sexual structures such as the peacock's tail. As Arnqvist concludes, the fact that male genitals are under sexual selection certainly blurs the traditional dichotomy¹¹ between primary reproductive organs — those that deliver ejaculates — and the secondary sexual structures of males such as fancy tails. In considering the intricate genital appendages of male insects, Darwin had a similar take on this issue when he

concluded that “it is scarcely possible to decide which ought to be called primary and which secondary”.

Darryl T. Gwynne is in the Department of Zoology, University of Toronto in Mississauga, Mississauga, Ontario, Canada L5L 1C6.

e-mail: dgwynne@credit.erin.utoronto.ca

1. Darwin, C. *The Descent of Man and Selection in Relation to Sex* 2nd edn (John Murray, London, 1874).
2. Eberhard, W. G. *Sexual Selection and Animal Genitalia* (Harvard Univ. Press, Cambridge, MA, 1985).
3. Arnqvist, G. *Nature* **393**, 784–786 (1998).
4. Proctor, H. C., Baker, R. L. & Gwynne, D. T. *Can. J. Zool.* **73**, 2010–2020 (1995).
5. Ridley, M. *Biol. Rev.* **63**, 509–549 (1988).
6. Ridley, M. *Anim. Behav.* **37**, 535–545 (1989).
7. Albuquerque, G. S., Tauber, C. A. & Tauber, M. J. *Evolution* **50**, 1598–1606 (1996).
8. Price, C. S. *Nature* **388**, 663–666 (1997).
9. Howard, D. J. & Gregory, P. G. *Phil. Trans. R. Soc. Lond. B* **340**, 231–236 (1993).
10. Eberhard, W. G. *Am. Nat.* **142**, 564–571 (1993).
11. Hunter, J. *Observations on Certain Parts of the Animal Oeconomy* (London, 1786).
12. Helfer, J. R. *How to Know the Grasshoppers, Crickets, Cockroaches and Their Allies* (Dover, New York, 1987).
13. Snodgrass, R. E. *The Male Genitalia of Orthopteroid Insects* Smithsonian Misc. Collections **96**, 1–107 (1937).

Chemical kinetics

Smaller, faster chemistry

Klavs Jensen

Chemical kinetics is the study of mechanisms and rates of chemical processes, in anything from biological systems to man-made materials. To measure the rate of reaction between two chemical species, they must be mixed much faster than the reaction rate — otherwise only the mixing rate is observed. Rapid enough mixing is difficult, especially for the many important liquid-phase and biological reactions with millisecond or shorter-timescale dynamics. But the device built by Knight *et al.*¹ (reported in *Physical Review Letters*) reduces the mixing timescale to less than 10 μ s, by vastly reducing the size of the reaction vessel.

Two reactants in a quiescent liquid mix by diffusion. The timescale is $\tau = L^2/D$, where L is the length scale over which diffusion must

take place and D is the diffusion coefficient. With typical values of D being 10^{-9} m² s⁻¹ or less for large molecules, mixing by diffusion is slow. Consequently, investigations of fast kinetics have traditionally speeded mixing by encouraging turbulence, which increases the contact area between two fluids and so reduces the diffusion distance. Turbulence can be created by mechanical stirring, forcing fluids through nozzles, or using the wake behind a ball. But these methods typically require high flow rates, and correspondingly large samples, so they are unattractive when dealing with expensive biological samples or chemicals that could damage the environment. Moreover, the mixed reactants must usually be pumped from such mixing systems to the diagnostic equipment, which introduces dead time, further limiting the

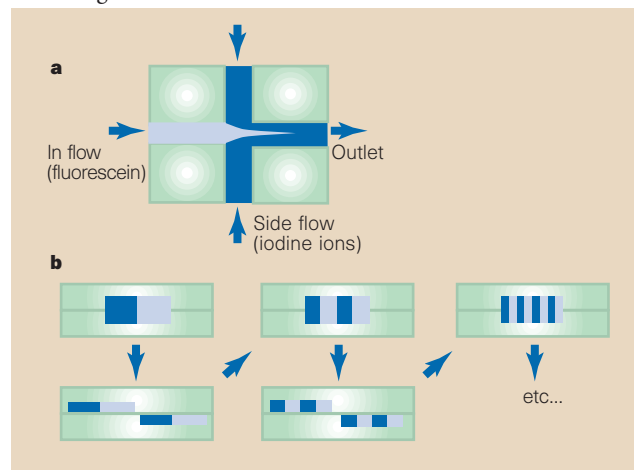


Figure 1 Microfluidic approaches to achieving fast mixing in laminar flows by diffusion. a, Hydrodynamic focusing¹, which works on the smallest scales, and thus the shortest timescales, yet achieved. b, Repeated layering of fluid elements in static mixer² (we see successive slices through channels pointing towards us).

minimum observable reaction timescale.

Knight and colleagues¹ have used a different approach, building a continuous-flow mixer in which the diffusion length L is reduced to micrometres. Diffusive mixing times of less than 10 μ s can then be achieved without having to induce turbulence. The mixer is made out of silicon, using the photolithography and Cl_2 reactive-etching processes normally used in integrated circuit manufacture. It has four 10- μ m-deep channels in a cross pattern, with each channel tapering down in width towards the common intersection (Fig. 1a). The side flows squeeze, or focus, the inlet flow into a thin stream, as thin as 50 nm, which makes diffusive mixing very rapid.

The chip is sealed on top with a transparent slip, allowing the mixing processes to be observed using a confocal scanning microscope. The inlet flow is labelled with fluorescein, whose fluorescence is quenched in a diffusion-controlled reaction with iodide ions coming in at the side channels. A simple electrical-circuit model explains the data — and can identify device modifications to achieve even faster mixing, such as increasing the ratio of inlet resistance to side-channel resistance.

Several different microfluidic approaches to achieving short diffusion lengths have been demonstrated. Most of the promising microstructures use successive lamination of the two streams to be mixed (Fig. 1b)^{2,3}. In this 'static mixer' scheme, the two fluids are brought into contact and the combined stream is separated into two flows along a line perpendicular to the mixing interface. These two streams are then recombined, doubling the fluid interface and halving the diffusion length. By repeating the process several times, very short diffusion distances are obtained, thereby producing rapid mixing. These mixers typically handle larger fluid volumes than the hydrodynamic focusing approach of Knight *et al.*, but at the expense of increased mixing times. Ultimately, the design of a micromixing unit is a trade-off between mixing speed, pressure drop, volume flow, ease of fabrication, and integration with chemical detection devices.

The new work demonstrates, in particular, the potential of a combination of micro-fabrication and modern microscopy for probing fast chemical reactions using very small sample sizes (microlitres to nanolitres). It is just one example of the rapidly developing applications of microfluidics to chemical and biological systems⁴. Micro-fabrication technologies, originally developed for semiconductor processing and extended to micro-electromechanical systems⁵, are rapidly being employed in chemistry and biology — for such uses as DNA synthesis and detection⁶, microchemical analysis⁷, and microchemical reactors for on-demand synthesis⁸.

Mixing is one of several operations, along with fluid transport, chemical reaction, product separation and chemical identification, that must be understood and optimized to fully realize the opportunities offered by these tiny chemical systems. The reduction in size makes mixing faster, but also increases the importance of surface effects such as surface tension. In the future, the combination of microfabrication techniques with new materials and chemical self-assembly approaches^{9,10} will allow chemical tailoring of surfaces, and so lead to further advances in microfluidic biological and chemical systems. $\square$

Klavs Jensen is in the Departments of Chemical Engineering and of Materials Science and Engineering, Massachusetts Institute of Technology,

77 Massachusetts Avenue, Cambridge, Massachusetts 02139-4307, USA.
e-mail: kfjensen@mit.edu

1. Knight, J. B., Vishwanath, A., Brody, J. P. & Austin, R. H. *Phys. Rev. Lett.* **80**, 3863–3866 (1998).
2. Möbius, H., Ehrfelt, W., Hessel, V. & Richter, Th. *Proc. Transducers* **95**, 775–778 (1995).
3. Branebjerg, J., Gravesen, P., Krog, J. P. & Nielsen, C. R. *9th Annu. Int. Workshop Micro Electro Mechanical Systems* (eds Allen, M. G. & Reed, M. L.) 441–446 (IEEE, New York, 1996).
4. Ehrfelt, W. (ed.) *Microreaction Technology* (Springer, Berlin, 1998).
5. Wise, K. D. & Najafi, K. *Science* **254**, 1335–1342 (1991).
6. Kopp, M. U., de Mello, A. J. & Manz, A. *Science* **280**, 1046–1048 (1998).
7. Harrison, D. J. *et al. Science* **261**, 895 (1993).
8. Srinivasan, R. *et al. Am. Inst. Chem. Eng. J.* **43**, 3059–3069 (1997).
9. Terfort, A., Bowden, N., Carbeck, J. & Whitesides, G. M. *Nature* **386**, 162–164 (1998).
10. Trau, M. *et al. Nature* **390**, 674–676 (1997).

Developmental biology

SMAD proteins and mammalian anatomy

Rik Derynck

How does the mammalian body plan, with its different anatomical structures and varied tissues, arise during development? The patterns and anatomical changes that occur have been well described, but the cell-made factors and signalling mechanisms that drive these changes have been more elusive. A paper by Nomura and Li¹ on page 786 of this issue and reports in *Cell*² and *Genes and Development*³ describe some of these factors and show that the signal transducers Smad2 and Smad4 are vital in early development.

Members of the transforming growth factor- β (TGF- β) superfamily regulate many important developmental processes, such as mesenchymal differentiation, skeletal morphogenesis and skin formation. The roles of most of these secreted proteins are poorly characterized, but others — such as the activins and several 'bone morphogenetic proteins' (BMPs) — are known to regulate specific events in both early and late development. Exactly how these factors act remained unclear until three years ago, when SMADs were discovered as intracellular proteins that mediate the effects of signalling from extracellular TGF- β -related factors (Fig. 1, overleaf). Now Nomura and Li¹, Waldrip *et al.*² and Sirard *et al.*³ describe the developmental consequences of targeted inactivation of the genes encoding the Smad2 or Smad4 proteins.

Gastrulation is an early developmental process in which the three primary layers of the animal body (ectoderm or outer layer, endoderm or inner layer, and mesoderm) are formed and positioned for further development. In mice, gastrulation starts with the formation of the primitive streak — a localized thickening of the epiblast (embryonic

ectoderm) occurs and marks what is to be the posterior end of the mouse embryo. As gastrulation proceeds, the primitive streak extends anteriorly and epiblast cells move through the primitive streak to form mesoderm. Expression of both the transcription factor Brachyury and the TGF- β -related protein nodal marks the cells of the primitive streak and the onset of mesoderm formation. The mesodermal cells then develop into several extra-embryonic and embryonic structures (Fig. 2).

Studies of the phenotypes of Smad2-defective mice^{1,2} now reveal that Smad2 signalling is essential for embryonic mesoderm formation and the establishment of anterior–posterior polarity. Nomura and Li¹ report the absence of mesoderm in Smad2-deficient embryos, and Waldrip *et al.*² see extra-embryonic mesoderm without formation of embryonic mesoderm, suggesting a less severe phenotype. These differences in phenotype may be due to differences in genetic background or in the strategy used by the authors to target the *Smad2* gene for deletion. Smad2 is also required to establish the anterior–posterior polarity, probably because of its function in restricting the site of primitive-streak formation. So Smad2 signalling is one of the earliest-known events required for the formation of mesoderm and anterior–posterior patterning.

The essential role of SMAD signalling in early development is also shown by the gastrulation defect of Smad4-deficient embryos, reported by Sirard *et al.*³. Although these embryos also lack mesoderm, they are morphologically different from Smad2-deficient embryos and have highly abnormal visceral endoderm development (the position of the visceral endoderm is shown in Fig.

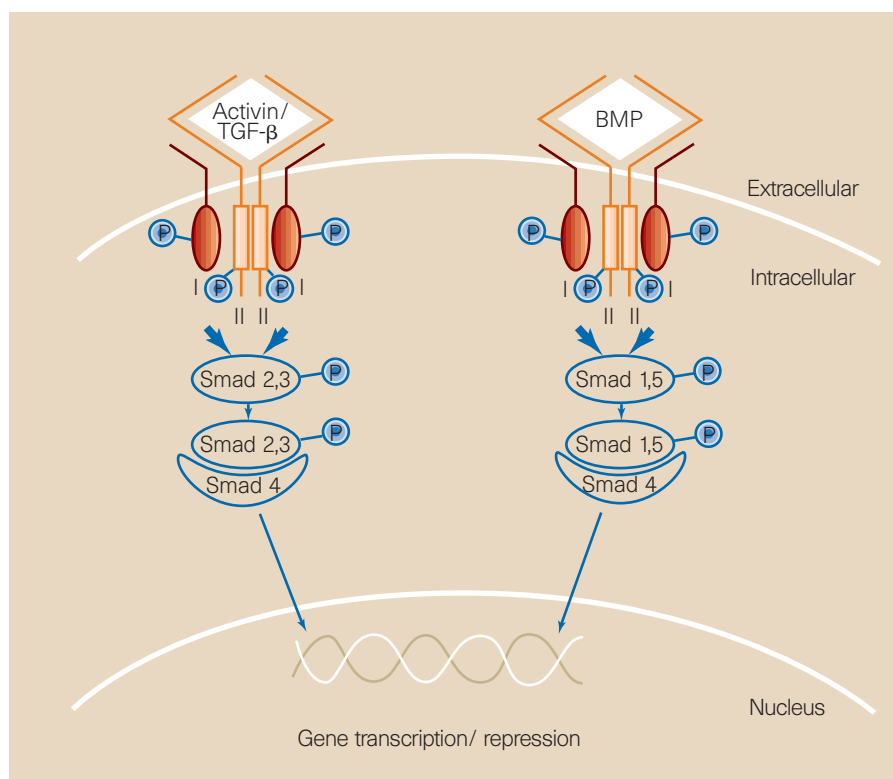


Figure 1 Signalling through SMAD proteins. Receptor complexes for transforming growth factor- β (TGF- β) and related factors, such as activins and bone morphogenetic proteins (BMPs), consist of two transmembrane serine/threonine kinases, the type II and type I receptors. Binding of activin, TGF- β or BMPs to the receptor complex results in phosphorylation of type I receptors by type II receptors. The activated type I receptors in turn phosphorylate, and thereby activate, a select set of SMADs: Smads 1 and 5 are activated by BMP receptors, whereas Smads 2 and 3 are phosphorylated by receptors for activin and TGF- β . These SMADs then dissociate from the receptors to form complexes with Smad4/DPC4, which is a putative tumour-suppressor protein. Smad4 is essential for signalling by TGF- β -related factors. These SMAD complexes move into the nucleus, where they may activate, or repress, transcription of different, defined genes¹⁴. Smads 6 and 7 (not shown) inhibit signalling, probably by competing for receptor binding and/or for association with Smad4/DPC4 (ref. 14). Phosphorylation is represented by a circled 'P'.

2). This phenotype may be due to a vital need for Smad4 expression at an earlier stage than Smad2, possibly to allow BMP-induced SMAD signalling during early gastrulation⁴.

The early defects and lack of mesoderm formation in the absence of Smad2 raise two related questions. First, why is this phenotype more severe than the phenotypes resulting from inactivation of TGF- β -related factors that activate Smad2? There are more than 40 TGF- β -related factors, but only five SMADs (Smads 1, 2, 3, 5 and 9) are activated by the ligand-stimulated receptors. This means that many TGF- β -related factors may signal through the same SMADs. Smad2 is activated by two activins and three TGF- β s, and probably by several other TGF- β -like factors, so inactivation of Smad2 may abol-

ish signalling from a set of TGF- β -related factors.

Second, which TGF- β -related ligand(s) induce Smad2 activation before mesoderm formation? Although activins have been implicated in the formation of mesoderm and activate Smad2, no gross defects in mesoderm formation occur in the absence of activins A and B (ref. 5). In contrast, inactivation of BMP-4 (ref. 4) or nodal⁶ results in severe early defects and impaired mesoderm formation. BMP-4 signals through Smads 1 and/or 5, and not through Smad2, so is unlikely to be the ligand that activates Smad2. This brings us to a possible connection between nodal and Smad2. Nodal is expressed before formation of the primitive streak, and its inactivation results in defects in this process⁶. The receptors and SMADs that mediate nodal activity are not known. Embryos lacking Smad2 or nodal have similar defects, so Nomura and Li¹ analysed the phenotypes of mice lacking one copy of the *Smad2* gene and one copy of the *nodal* gene. Their findings indicate that nodal may signal through Smad2, although biochemical analyses are required to test this possibility. Other ligands, possibly of maternal origin, are also likely to be required for the Smad2 activation that leads to the formation of mesoderm.

Nomura and Li¹ also found that developmental changes depend on the amount of Smad2 activity. A defective phenotype is usually only apparent when both of a pair of genes are inactivated. But the presence of only one functional *Smad2* gene, and the resulting lower-than-normal levels of Smad2 protein, also results in developmental

defects, albeit much less severe than those seen when both genes are inactivated. This striking dosage-dependence of Smad2 signalling — apparent both in early development and at later stages — is consistent with the observation that, in the African clawed toad *Xenopus laevis*, different mesodermal markers and cell types are induced by different concentrations of activin⁷, nodal⁸ or Smad2 (ref. 9). Different thresholds for transcriptional activation of individual genes by Smad2, and competition of Smad2 with other SMADs for binding to gene promoters, are likely to be the basis for these dosage-dependent changes.

Finally, the new results¹ highlight the role of TGF- β -related factors in defining left-right asymmetry. Visceral organs and the position and organization of the heart are normally asymmetric. Little is known about how this left-right patterning occurs, but TGF- β -related factors such as nodal^{10,11}, lefty¹² and cNR-1 (ref. 13) are asymmetrically expressed and may regulate left-right patterning. SMAD signalling may relay information from these ligands; consistent with this idea, decreased Smad2 signalling resulted in various abnormalities that reflect disturbances in the left-right asymmetric pattern¹.

The inactivation of Smads 2 and 4 has made it possible to study how this class of signalling effectors is involved in establishing the body plan and in driving many developmental changes. The results emphasize the important functions of TGF- β -related factors at many stages in development, from the fertilized egg to the adult organism. $\square$ Rik Derynck is in the Departments of Growth and Development, and of Anatomy, Programs in Cell

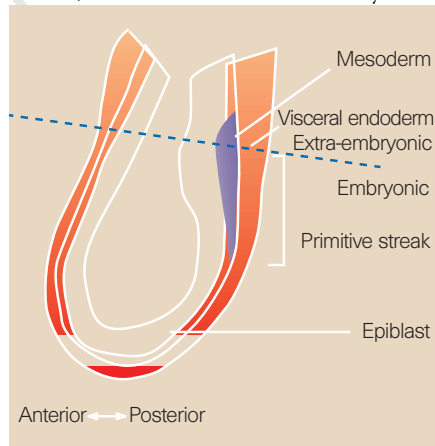


Figure 2 Representation of a mouse embryo at early gastrulation (embryonic day of development 6.5), indicating the different cell layers. The site of mesoderm formation is shown in purple.

Biology and Developmental Biology, University of California at San Francisco, San Francisco, California 94143-0640, USA.

e-mail: derynck@itsa.ucsf.edu

1. Nomura, M. & Li, E. *Nature* **393**, 786–790 (1998).
2. Waldrup, W. R. *et al.* *Cell* **92**, 797–808 (1998).
3. Sirard, C. *et al.* *Genes Dev.* **12**, 107–119 (1998).
4. Winnier, G. M., Blessing, M., Labosky, P. A. & Hogan, B. L. M. *Genes Dev.* **9**, 2105–2116 (1995).
5. Matzuk, M. M. *et al.* *Nature* **374**, 354–356 (1995).
6. Conlon, F. L. *et al.* *Development* **120**, 1919–1928 (1994).
7. Green, J. B., New, H. V. & Smith, J. C. *Cell* **71**, 731–739 (1992).
8. Jones, C. M. *et al.* *Development* **121**, 3651–3662 (1995).
9. Baker, J. C. & Harland, R. M. *Genes Dev.* **10**, 1880–1889 (1996).
10. Collignon, J., Varlet, I. & Robertson, E. J. *Nature* **381**, 155–158 (1996).
11. Lowe, L. A. *et al.* *Nature* **381**, 158–161 (1996).
12. Meno, C. *et al.* *Genes to Cells* **2**, 513–524 (1997).
13. Levin, M. *et al.* *Cell* **82**, 803–814 (1995).
14. Heldin, C.-H., Miyazono, K. & ten Dijke, P. *Nature* **390**, 465–471 (1997).

Signal transduction

Actin, cofilin and cognition

Jody Rosenblatt and Timothy J. Mitchison

Patients with Williams syndrome, a multi-gene deletion syndrome, suffer from mild mental retardation and vascular disease. They also have defects in visuo-spatial cognition — a failure to integrate parts into a whole — that are linked to deletion of the gene that codes for LIM-kinase 1 (LIMK-1)¹. Consistent with this cognitive defect, large amounts of LIMK-1 are expressed in neurons², yet its molecular targets have remained elusive until now. On pages 805 and 809 of this issue, however, Arber *et al.*³ and Yang *et al.*⁴ report that LIMK-1 phosphorylates cofilin, an essential protein that is required for turnover of actin filaments.

During cell movements such as neuron outgrowth or leukocyte chemotaxis, actin filaments must be organized into a dense, dynamic meshwork. This forms at the leading edge of a cell, where actin polymerization drives forward movement, and it usually takes the form of thin sheets (lamellipodia) or spikes (filopodia). Cell movement is a dynamic process, and actin at the leading edge of the cell must be continuously depolymerized and then repolymerized to produce this movement⁵. Actin depolymerization limits the length of lamellipodia and enables the actin subunits to be recycled for further rounds of polymerization.

There is mounting evidence that the key enzyme required for actin depolymerization is cofilin. *In vivo*, cofilin has been shown to be essential for cytokinesis⁶, endocytosis⁷ and other cell processes that require rapid turnover of actin filaments⁸. *In vitro*, cofilin binds to both actin monomers and polymers, and promotes the disassembly of actin filaments. Cofilin is regulated by phosphorylation of the serine residue at position 3, which inhibits its actin-binding and depolymerization activities. Stimuli that induce the production of lamellipodia relieve this inhibition by causing the rapid dephosphorylation of cofilin⁹.

Arber *et al.*³ and Yang *et al.*⁴ now provide evidence that LIMK-1 phosphorylates (and therefore inactivates) cofilin. Both groups labelled mammalian cells with radioactive inorganic phosphate, and found that isolated LIMK-1 associates with only one phos-

phoprotein — cofilin. Moreover, LIMK-1, but not an inactive form of the enzyme, can phosphorylate recombinant cofilin. These findings account for the observations that overexpression of LIMK-1 leads to accumulation of excess actin filaments, whereas overexpression of a dominant-negative LIMK-1 (a mutant form that disrupts the wild-type activity) inhibits the accumulation of actin filaments.

For cells to move, signals from their peripheries must be relayed to proteins such as LIMK-1 and cofilin. What factors relay these signals? Arber *et al.* and Yang *et al.* show that the small GTPase Rac may be important in regulating the activity of LIMK-1. Rac regulates the actin reorganization that is required to form lamellipodia¹⁰, yet few of its protein targets have been identified. The authors^{3,4} found that Rac-dependent formation of lamellipodia is blocked by dominant-negative forms of LIMK-1, suggesting that LIMK-1 acts downstream in the Rac pathway. Moreover, dominant-negative Rac leads to decreased phosphorylation of cofilin, whereas activated Rac modestly increases phosphorylation. These results indicate that Rac activates LIMK-1 which, in turn, phosphorylates — and inactivates — cofilin (Fig. 1, overleaf).

To induce the formation of lamellipodia, Rac must do more than simply inactivate cofilin — it must also induce the polymerization of actin. One possible mechanism involves Rac-induced increases in the levels of phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P₂), which is thought to cause filament uncapping¹¹. Removal of a capping protein from the end of an actin filament could then allow the polymerization of actin to resume. Actin also needs to be depolymerized for the formation of lamellipodia, so we might expect that the Rac pathway only transiently inactivates cofilin. Indeed, Yang *et al.*⁴ find no net change in cofilin phosphorylation when endogenous (not overexpressed) Rac is activated. Transient inactivation of cofilin at the leading edge could allow nearby actin filaments to grow. Additionally, cofilin phosphorylation may induce the release of recently depolymerized actin monomers. Continuous cycles of cofilin phosphoryla-



100 YEARS AGO

All persons who interest themselves in the progress of celestial mechanics, but can only follow it in a general way, must feel surprised at the number of times demonstrations of the stability of the solar system have been made. Lagrange was the first to establish it, Poisson then gave a new proof; afterwards other demonstrations came, and others will still come. Were the old demonstrations insufficient, or are the new ones necessary? The astonishment of these persons would doubtless be increased if they were told that perhaps some day a mathematician would show by rigorous reasoning that the planetary system is unstable. ... It can be shown, in certain particular cases, that the elements of the orbit of one planet will return an infinite number of times to very nearly the initial elements, and that is also probably true in the general case; but it does not suffice. It should be shown that these elements will not only regain their original values, but that they will never deviate much from them. This last demonstration has never been given in a definite manner, and it is even probable that the proposition is not strictly true. — M. H. Poincaré.

From *Nature* 23 June 1898.

50 YEARS AGO

The French Association of Scientific Workers is organising an International Holiday Centre at Morzine (Haute-Savoie) during July 15–September 1. Morzine, without being a fashionable holiday resort, is one of the beauty spots of the French Alps. It lies to the south of Thonon at a height of about 1,000 metres; and around it are ranges rising to about 2,500 metres. So far as hotels are concerned, the French Association of Scientific Workers can secure accommodation at 800 and 700 francs a day (including service and taxes). It has also secured some good camping sites, and is trying to make arrangements for a canteen. Bookings can only be made for whole weeks (from Monday morning to Sunday night) and for a minimum of two weeks; registration fee, 500 francs each person. Applications should be made as soon as possible to the Association des Travailleurs Scientifiques, Maison de l'Université Française, 47 Boulevard St. Michel, Paris 5.

From *Nature* 26 June 1948.

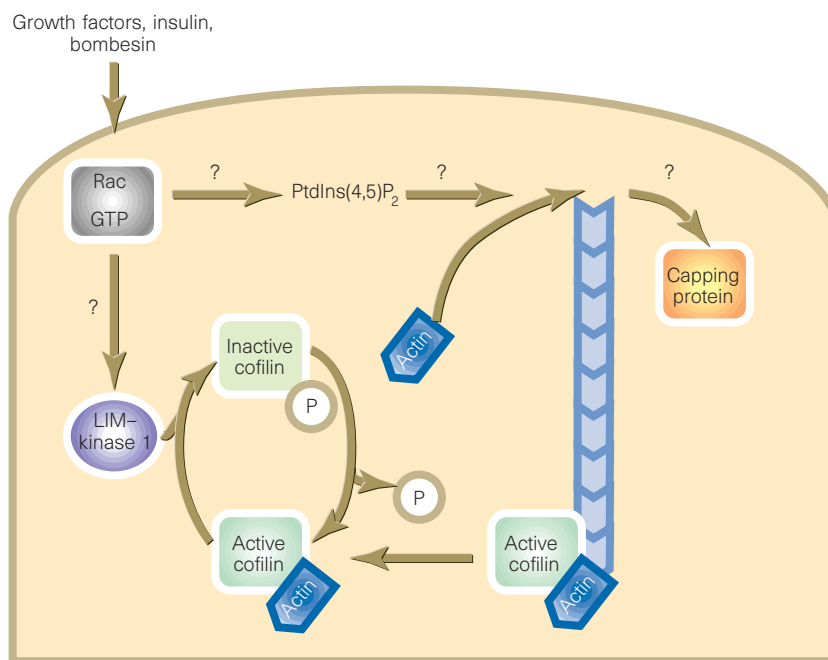


Figure 1 Model for Rac-induced changes in actin dynamics to form a lamellipodium, based on the findings of Arber *et al.*³ and Yang *et al.*⁴. When Rac is stimulated by growth factors, it activates LIM-kinase 1 (LIMK-1) which, in turn, phosphorylates cofilin. This may inhibit localized disassembly of actin filaments and help to release cofilin from the actin monomers. Rac-mediated stimulation of actin polymerization may occur through uncapping or increased nucleation of actin filaments. Both of these mechanisms may occur through production of phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P₂) by phosphatidylinositol phosphate 5-kinase. The intermediates that activate phosphatidylinositol phosphate 5-kinase and LIMK-1, and the factor that stimulates dephosphorylation of cofilin, are unknown.

tion and dephosphorylation would allow both cofilin and actin to be recycled for further rounds of depolymerization and polymerization, respectively. Clearly, further work is needed to sort out the role of cofilin phosphorylation in actin dynamics, and to clarify the temporal and spatial regulation of actin depolymerization in cells.

Another exciting aspect of this work comes from studies on patients with Williams syndrome, who have only one copy of the *LIMK-1* gene. Given that LIMK-1 regulates the turnover of actin filaments, why should people with Williams syndrome have defects in visuo-spatial cognition, as opposed to other processes that require actin dynamics? Perhaps neurons require high levels of LIMK-1 to finely regulate the turnover of actin filaments during axonal guidance. The PC12 neuronal cells containing dominant-negative LIMK-1 studied by Arber *et al.* may provide a clue as to what neurons with decreased levels of LIMK-1 look like. Although neuron outgrowth still occurred, the neurites contained dramatically fewer filopodia — finger-like projections that are thought to sense in which direction axons should grow. If this is true, a decrease in LIMK-1 could account for the abnormally clustered and aligned neurons seen in the brains of patients with Williams syndrome¹².

To show that a decrease in LIMK-1 leads to abnormal neuronal wiring, we need to study people who lack a copy of the *LIMK-1* gene (and not contiguous genes), as well as mouse knockout models. Future work will also need to address whether the visuo-spatial cognitive defect in people with Williams syndrome results exclusively from phosphorylation of cofilin by LIMK-1, or whether other substrates for LIMK-1 exist. □

Jody Rosenblatt is in the Department of Biochemistry, University of California at San Francisco, 513 Parnassus Avenue, San Francisco, California 94143, USA.

e-mail: jody@cgl.ucsf.edu

Timothy J. Mitchison is in the Department of Cell Biology, Harvard Medical School, Room C517, 240 Longwood Avenue, Boston, Massachusetts 02115, USA.

e-mail: Timothy_Mitchison@hms.harvard.edu

1. Frangiskakis, J. M. *et al.* *Cell* **86**, 59–69 (1996).
2. Okano, I. *et al.* *J. Biol. Chem.* **270**, 31321–31330 (1995).
3. Arber, S. *et al.* *Nature* **393**, 805–809 (1998).
4. Yang, N. *et al.* *Nature* **393**, 809–812 (1998).
5. Theriot, J. A. & Mitchison, T. J. *Nature* **352**, 126–131 (1991).
6. Abe, H., Obinata, T., Minamide, L. S. & Bamburg, J. R. *J. Cell Biol.* **132**, 871–885 (1996).
7. Lappalainen, P. & Drubin, D. G. *Nature* **388**, 78–82 (1997).
8. Theriot, J. A. *J. Cell Biol.* **136**, 1165–1168 (1997).
9. Moon, A. & Drubin, D. G. *Mol. Biol. Cell* **6**, 1423–1431 (1995).
10. Ridley, A. J., Paterson, H. F., Johnston, C. L., Diekmann, D. & Hall, A. *Cell* **70**, 401–410 (1992).
11. Hartwig, J. H. *et al.* *Cell* **82**, 643–653 (1995).
12. Galaburda, A. M., Wang, P. P., Bellugi, U. & Rossen, M. *Neuroreport* **5**, 753–757 (1994).

Daedalus

Muffled furnace

Noise is one of the major nuisances of modern life. Yet the usual sound-absorbing materials are purely passive, and can never damp it out completely. Daedalus now proposes an active absorber, inspired by the observation in a chemical textbook that the carbon monoxide flame “gives a curious impression of silence”. Carbon monoxide, he notes, burns with a reduction in the number of gas molecules. If, like most gas reactions, the burning speeds up with pressure, then a sonic pressure-peak will deplete the gas of molecules at an enhanced rate, and damp itself out. Conversely, when the pressure is low, the depletion slows down. So the flame absorbs the sound. Furthermore, certain flames are extremely sensitive to sound. They were used as acoustic detectors in pre-microphone days. Even a weak sound changed their combustion regime very visibly.

So Daedalus is inventing quiet flame technology. He is devising nozzles and flame-surrounds that optimize this sound-damping effect. He hopes to perfect a gas burner whose nonlinear reaction regime overreacts to sound, and thus absorbs it perfectly. A street lamp that absorbed traffic noise would be welcome on busy roads; victims of aircraft noise or pop-crazed neighbours would love a gas fire that gave out quiet as well as heat.

Even so, nobody would want to keep a fire burning on a hot day, merely to keep the noise down. So Daedalus is taking the argument further. When iron rusts, for example, gas molecules are completely absorbed into a solid. DREADCO's chemists are now studying the oxidation of iron alloys, as well as yellow phosphorus, aluminium amalgam and even lithium (which absorbs nitrogen as well). They are seeking pressure-sensitive reaction regimes with strong nonlinearity, or even a pressure threshold. Their goal is a surface that rusts or tarnishes with total absorption of sound. Ideally it should change colour during the reaction. When fully reacted, it could be regenerated, perhaps by reduction with hydrogen.

This novel decor will be very expensive at first, and will be aimed at acoustics laboratories and recording studios. Gradually it should spread to the more opulent homes, offices and public buildings. A personal version in earmuff form would be widely welcomed; not only for the quiet in which it wrapped the wearer, but for the pleasing warmth released by its slow sonic oxidation.

David Jones

The end of the old model Universe

Peter Coles

The 'Big Bang' model includes a few unknown numbers that determine the size, shape and future of the Universe. In the past few years our measurements of these free parameters have begun to rule out the old picture of a Universe dominated by cold dark matter. What will take its place?

“Exploding stars”, “ripples at the edge of space”, “antigravity”, a “Universe the size of a pea”. All of these have made somewhat garbled appearances in the press in the past few months. But there is more to these announcements than hype — they are the public manifestation of dramatic developments in cosmology that have fundamentally changed the character of the subject. Observations only recently made possible by improvements in astronomical instrumentation have put theoretical models of the Universe under intense pressure. The standard ideas of the 1980s about the shape and history of the Universe have now been abandoned — and cosmologists are now taking seriously the possibility that the Universe is pervaded by some sort of vacuum energy, whose origin is not at all understood.

The ‘Big Bang’ model, in essence, stems from the observation that as the Universe is expanding, it must have been smaller in the past, and also denser and hotter. But in its modern form, the Big Bang model is based on mathematical solutions of the equations

of General Relativity, originally obtained by Friedmann in the 1920s. These describe the dynamics and geometry of a Universe obeying the cosmological principle — the principle that, on large enough scales, the Universe is the same in every place (homogeneous) and looks the same in every direction (isotropic).

Weakness

Each solution (Box 1, overleaf) is described by two parameters: the Hubble constant H_0 and the density parameter Ω_0 . An even more general class of models can be obtained by throwing in a cosmological constant (usually known as Λ ; Box 2, page 744). These three parameters together determine the size, age, geometry and expansion rate of the Universe. But the model has a weakness: the numerical values of these parameters are not predicted by theory because physical law breaks down at the origin of time, and so the initial conditions on which all the subsequent evolution depends are not known. The numbers must instead be inferred from observations. That is why the Big Bang is bet-

ter described as a model than as a theory: its predictive power is restricted by the presence of the free parameters.

The Big Bang model accounts for three basic but important cosmological observations¹: the expansion of the Universe (specifically, the Hubble law, describing the linear dependence of recession velocity on distance from the observer); the existence of the 3 K microwave emission from all directions on the sky (the cosmic microwave background, or CMB, which comes from the time when the Universe became transparent); and the abundances of light atomic nuclei, such as helium and deuterium, which were created in the first few minutes of the Universe. But such is the uncertainty introduced by the irksome unknown parameters that testing the Big Bang further has been like trying to hit a moving target.

The second weakness of the Friedmann Big Bang model is that it describes a highly idealized cosmos that is perfectly homogeneous and isotropic — a far cry from the lumpy, irregular beast we observe around us. If the theorists’ model is like milk, then the real Universe is more like porridge. The Big Bang does not deserve to be called a theory unless and until it can explain how the lumps in the porridge, galaxies and clusters of galaxies, came into being and evolved.

The most widely accepted picture of how structure formed involves the idea of gravitational instability. A perfectly smooth self-gravitating fluid with the same density everywhere stays homogeneous for all time. But any slight irregularities (which always exist in reality) tend to get amplified by the action of gravity. A small patch of the Universe that is slightly denser than average tends to attract material from around itself; it therefore gets even denser and attracts even more material. This instability will form a highly concentrated lump, held together

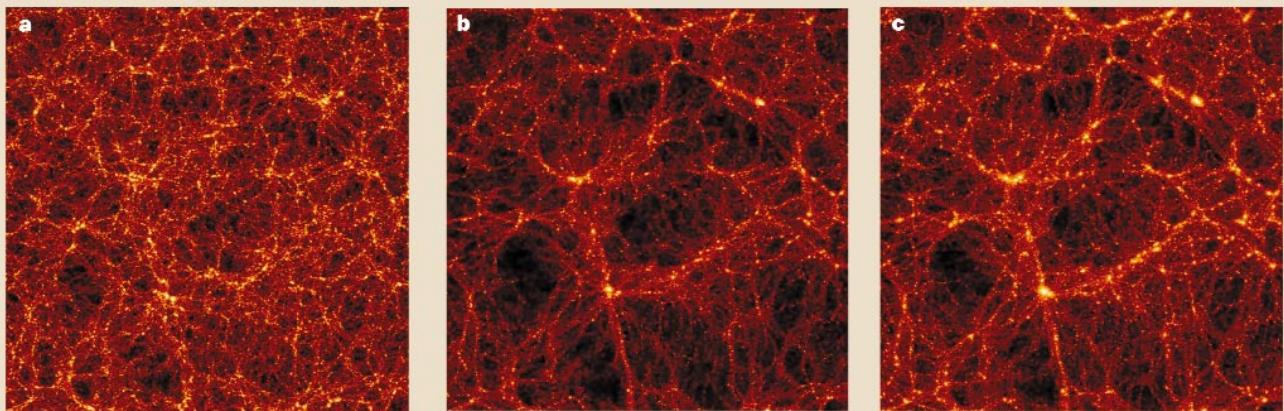


Figure 1 Playing with the Universe: simulations of how galaxies cluster together to form larger structures, with different values of the cosmological parameters (Box 1). a, In the standard $\Omega_0 = 1$ Universe, dominated by cold dark matter, the largest structures (‘superclusters’) are not as large as those seen in real galaxy surveys. Adding some ‘hot’ dark matter (such as massive neutrinos) only smears out structure further. b, For a less dense, open Universe with $\Omega_0 = 0.3$, the structures are larger. But this model has problems coping with observed microwave background fluctuations. c, A cosmological constant (Box 2) is added, to make the ‘new model Universe’ with $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$. This fits all observations so far.

THE VIRGO COLLABORATION

Box1 Cosmology by numbers – a primer

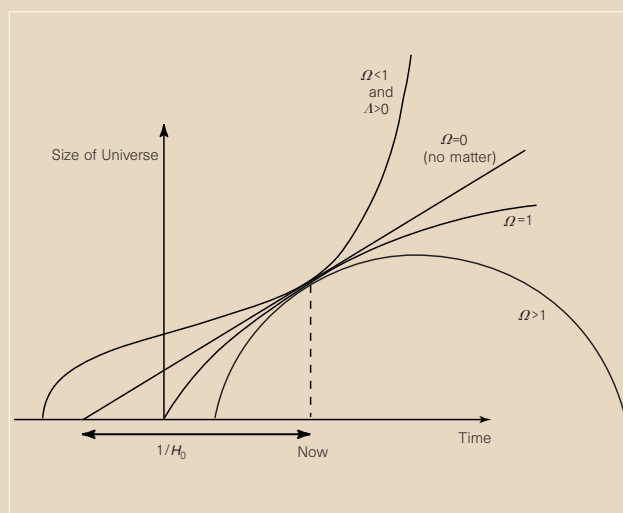
The equations describing cosmology are complicated, but the essentials can be encapsulated in just three numbers – numbers that can, in principle, be determined by observation.

The most famous of these parameters is the Hubble constant, H_0 . This appears in the Hubble law, $v = H_0 d$, relating the apparent recession velocity of a galaxy, v , to its distance from the observer, d . So the Hubble constant fixes the present expansion rate of the Universe.

But the Universe has not always been expanding at the same rate, nor will it do so in the future, because gravity decelerates it. If matter is dense enough, it will eventually halt the expansion of the Universe, and cause it to recollapse. If not, the Universe will expand forever. These two models are distinguished by the second cosmological parameter, Ω_0 . This is the ratio of the actual density of matter in the Universe to the

critical value, the dividing line between eternal expansion and ultimate recollapse. $\Omega_0 = 1$ is the critical value, so $\Omega_0 < 1$ means an ever-expanding Universe and $\Omega_0 > 1$ indicates one that eventually recollapses in a Big Crunch (as shown in the figure here). Whatever the value of Ω_0 , however, the effect of matter is always to slow down the expansion of the Universe.

This parameter also determines the curvature of space. In models with $\Omega_0 < 1$, space has negative curvature, like a hyperboloid or saddle. This is called an 'open' Universe. Space is positively curved ('closed') like a sphere if $\Omega_0 > 1$. In between, there is the compromise Universe, poised between eternal expansion and eventual recollapse, with $\Omega_0 = 1$. This is a Universe much favoured by theorists, with flat spatial geometry wherein Euclid's theorems apply. A closed Universe must be finite; flat and open Universes can be infinite or finite with appropriate



boundary conditions.

All this applies to models in which empty space has no inherent energy. But if it has, and there is a contribution to the density of the Universe from vacuum energy (measured by the 'cosmological constant', Λ ; see Box 2, page 744), then one can have a Universe in which the matter density is low but which is spatially flat. Moreover, a cosmological constant term, or something very like it¹⁴, is the only way

to produce a Universe that is not decelerating but accelerating.

These three parameters define the basic properties of the observable Universe, particularly its size and age. For a fixed value of the Hubble constant, models with low Ω_0 are larger and older than if $\Omega_0 = 1$. And with both $\Omega_0 < 1$ and a cosmological constant to make space flat, the Universe would be larger and older still. **P.C.**

by gravitational forces.

Gravitational instability is very rapid if the fluid one starts with is static, but much slower if it is expanding, because then gravity must overcome the outward motion of the fluid. So for structure to exist now, fluctuations must have been reasonably large in the very early Universe. Moreover, the rate of growth of these fluctuations depends very sensitively on the cosmological density parameter Ω_0 (Box 1), because the more matter there is, the faster the instability acts.

This would all be academic were it not possible to measure the fluctuations present in the early Universe, by measuring the variations in sky temperature of the cosmic microwave background. These correspond to fluctuations in the density of matter that were present about 300,000 years after the Big Bang (the present age of the Universe is around 12 billion years). It is therefore possible to trap the density parameter Ω_0 in a pincer movement, between measurements of galaxy clustering around us today and measurements of the temperature variations on the microwave sky.

Cosmologists knew all this as early as the 1970s, but attempts to detect temperature

variations in the microwave background yielded only upper limits: the sky was much smoother than the available instruments could gauge. But at this time the models of structure formation assumed a cosmos made only of ordinary 'baryonic' matter (mainly protons and neutrons); and such a fluid can have no more than about a twentieth of the critical density — much more, and the dense early Universe would have created more helium and less deuterium than we see². So baryonic models had to have low values of Ω_0 , and therefore predicted very high microwave background fluctuations, higher than the upper limits imposed by experiments. What was wrong?

A provocative solution appeared in the 1980s. It was realized that the overall density of the Universe might be much higher than the density of baryons if the extra material were some exotic substance that did not participate in nuclear reactions. The most favoured form of this postulated non-baryonic matter was dubbed 'cold dark matter', because it consisted of massive particles that had low thermal velocities when produced in the Big Bang. The cold dark matter (CDM) model of structure formation³ arrived on the

scene in the mid-1980s. This model was a theorists' dream. It assumed a flat Universe (Box 1) with $\Omega_0 = 1$ (mostly cold dark matter) and $\Lambda = 0$, the simplest of all the Friedmann solutions to work with. It also had a low Hubble constant, $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$, because higher values of H_0 with $\Omega_0 = 1$ produce very young Universes, younger than 13 billion years.

Inflation

The CDM particles were also easy to handle: assumed to be interacting only through gravity, their behaviour could be simulated straightforwardly on computers. And this picture also connected nicely with new ideas emerging from studies of the early Universe, including the new theory of cosmological inflation, which posits a very rapid expansion of the Universe when it is about 10^{-35} seconds old⁴. Inflation makes definite predictions about the form of the initial fluctuations⁵, and also, crucially, requires the Universe to be flat.

The CDM Universe appeared to be broadly consistent with the available data. All seemed well. But cracks started to appear in the early 1990s. Galaxy surveys were pub-

lished that mapped the local Universe out to larger scales than had previously been possible: and it appeared to be more lumpy on these scales than the CDM theory predicted. Then, in 1992, the Cosmic Microwave Background Explorer (COBE) satellite finally detected fluctuations on the microwave sky⁶. They were stronger than the CDM model predicted. As even more data have been analysed and digested, most cosmologists have come to conclusion that something is badly wrong with the CDM Universe.

Yet so attractive is the CDM idea that many workers have tried to retain the essential components of the model, adding extra ingredients to bring it into agreement with observations. So there arose a suite of variations on the basic theme, such as a CDM model with density less than a critical density, a CDM model with a cosmological constant, a CDM model with an admixture of a different component of 'hot' dark matter, and so on (Fig. 1). None of these has the persuasive simplicity of the original CDM model; they were assembled with the intention of fitting the data better, in the manner of Ptolemy's epicycles.

While these developments were taking place on the structure formation side, evidence was also being gathered by astronomers who were trying to nail down the cosmological parameters directly. The Hubble Space Telescope offered the chance to calibrate the extragalactic distance scale more accurately. Early results suggested an uncomfortably high value of H_0 , around 80

$\text{km s}^{-1} \text{Mpc}^{-1}$, corresponding to a small and young Universe, seemingly younger than the oldest stars — clearly an absurdity⁷. But later measurements, and revisions of the estimated ages of globular-cluster stars using results from the Hipparcos satellite⁸, seem to have resolved this problem. It now looks as though H_0 is 60–70 $\text{km s}^{-1} \text{Mpc}^{-1}$.

Density

Astronomers also kept trying to measure the density of matter, through its gravitational effects. The high orbital velocities seen around galaxies and within galaxy clusters confirmed the presence of dark matter in large quantities — corresponding to an overall density of around 30% of the critical density, or $\Omega_0 \approx 0.3$. Dynamical studies on larger scales, involving the gravitational effect of enormous superclusters, were more ambiguous, some indicating higher values of Ω_0 than this, some lower. But there was no clear evidence⁹ of the huge amounts of dark matter required if $\Omega_0 = 1$.

Direct evidence against the 1980s picture of $\Omega_0 = 1$ CDM has come in the past few years, with the resurrection and transformation of 'classical cosmology' — a battery of observational techniques designed to measure directly the geometry and deceleration rate of the Universe using observations of distant light sources¹⁰. In particular, two large-scale observational programmes are now devoted to detecting a particular kind of distant supernova explosion (called type Ia), which should act as 'standard candles'. If theoretical understanding of these events is

correct, then each one corresponds to a nuclear detonation of the same amount of material (a white dwarf star, topped up to a critical mass by gradually accreting material from a binary companion), and each explosion should therefore have the same power. Knowing the energy of the explosion is crucial because, having measured the received energy, one can then work out how much the Universe has expanded between the explosion time and the time that the radiation reached the telescope. Having different supernovae at different distances allows one to measure the deceleration rate of the Universe. These surveys indicate that the Universe is not decelerating quickly (not as quickly as if $\Omega_0 = 1$) — and it may even be accelerating^{11,12}.

Future

Where does all this leave us? How is the Universe changing? As far as structure formation is concerned, we know that the 'standard' CDM model is not the whole story. But the resurgence in classical cosmology has given us important clues as to precisely how this model might be modified to bring it into line with the Universe we see. The higher value of H_0 , the small value of Ω_0 inferred from dynamical arguments, and the weak deceleration seen in the supernova studies all suggest a global density of matter much less than the value required to make the Universe recollapse.

So, might we live in an open CDM Universe with $\Omega_0 \approx 0.3$ and $H_0 \approx 70 \text{ km s}^{-1} \text{Mpc}^{-1}$? At first sight, this seems an attractive model, as it does not require any exotic extras beyond the CDM particles. But there are reasons to be nervous about its prospects. For one thing, it sits uncomfortably with the idea of cosmological inflation, which, in its usual form, prefers a flat Universe. But then nobody knows whether inflation actually happened anyway. Some cosmologists have claimed that 'open' Universes (with negative spatial curvature; Box 1) can result from inflation, but that is controversial.

More seriously, this version of CDM has problems in matching observations of galaxy clustering with new data on cosmic microwave background fluctuations (Fig. 2). Useful though COBE was, its capabilities were limited. In particular, its angular resolution was only around 10 degrees, so COBE painted only a very broad-brush picture of the sky. More direct clues about the value of Ω_0 should be found in the structure of the temperature variations on scales much smaller than this, around one degree. There should be a characteristic pattern of hot spots and cold spots, whose size depends sensitively on Ω_0 , H_0 and Λ . Dozens of ground-based and balloon-borne experiments are in now progress that aim to detect these features, which are often called the

MILLIMETER ANISOTROPY EXPERIMENT

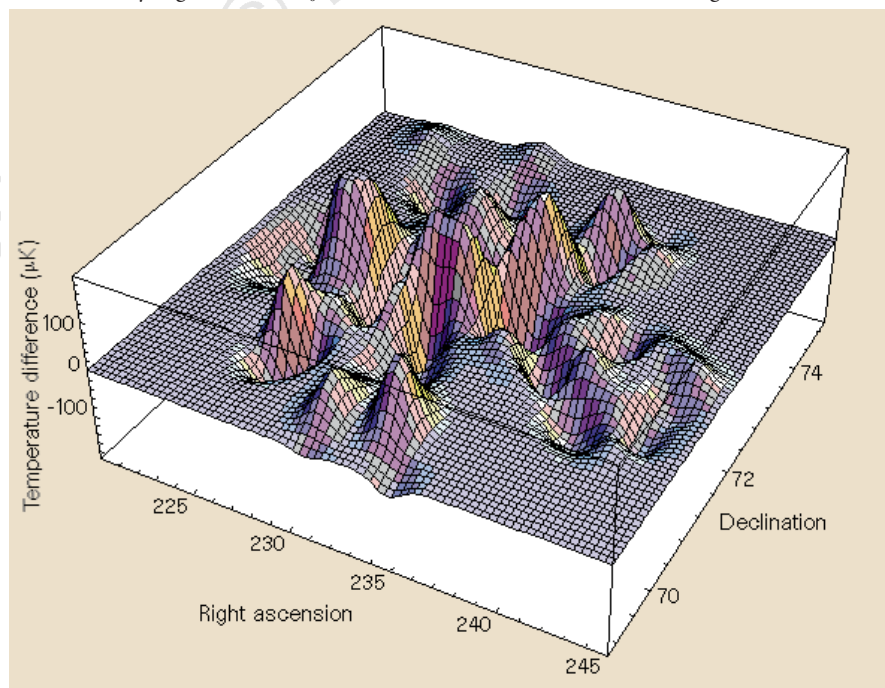


Figure 2 Microwave hot spots. The cosmic microwave background has an average temperature of 2.726 kelvin, but it fluctuates slightly in temperature, marking primordial fluctuations in the density of matter. Here is a small section of the sky, showing fluctuations on an angular scale of around a degree — fluctuations that seem to be too large for us to live in an open Universe with $\Omega_0 = 0.3$ and no cosmological constant.

Box2 The cosmological constant

The cosmological constant was introduced into the theory of General Relativity by Einstein, because he wanted to construct a cosmological model that was static in time – this was before Hubble's discovery of the expansion of the Universe. That made Einstein regret the addition of this term, but its legacy lingers.

The field equations of General Relativity can be written $G = \kappa T$, where G is the Einstein tensor and T the energy-momentum tensor, complicated matrices describing the time evolution and curvature of space-time (G) and the distribution and motion of matter and radiation (T). κ is a combination of fundamental constants.

Einstein added a new term, called Λ , to the left-hand side of the equation.

Effectively this modifies the law of gravity so that on large enough scales there is a new force. If Λ is positive it can be understood as a cosmic repulsion; if it has the opposite sign, as a cosmic attraction.

One might imagine that the cosmological constant would have vanished from the scene with the discovery of the expansion of the Universe, but that did not happen. Cosmologists began to ponder the consequences of interactions going on in the very early Universe. They theorized that the Universe could undergo phase transitions between states with different sets of fundamental forces (an example is the splitting of electromagnetic forces and the weak nuclear force from a unified 'electroweak' force). This gives rise to a

cosmological constant term, but one written on the right-hand side of Einstein's theory as part of T , leaving G (and the laws of gravity) undisturbed.

In this form, Λ acts like a kind of matter, but with very peculiar properties. It corresponds to a universal energy density that bends space in the same way as matter does; but it also has a negative pressure, and it is the gravitational effect of this pressure that causes the cosmic acceleration.

The energy density associated with the cosmological constant is not possessed by matter or radiation, but by 'empty' space. It acts like an extra contribution to Ω_0 of $\Omega_\Lambda = \Lambda/3H_0^2$. Each time space undergoes a phase transition, some amount of vacuum energy is expected

to be produced. Physicists working on the theory of elementary particles have tried to calculate the net amount of vacuum energy produced by all the phase transitions the Universe undergoes as it cools. The answer they get is catastrophically large: about 10^{120} times greater than the density of all the matter in the Universe¹⁵.

Such a result is at odds with observations, to put it mildly. If cosmological observations require the existence of a cosmological constant (and there are hints from supernova surveys that they might), the size of the term needed will correspond to a vacuum energy density similar to the density of matter, about 10^{120} times smaller than the predictions of particle physics.

P.C.

Doppler peaks. Existing maps appear to show them, but the data are rather noisy and the exact size of the peaks difficult to discern. They do appear, however, to be too large¹³ to have been produced in an open Universe with $\Omega_0 = 0.3$.

These observations, taken together with the possibility suggested by the supernova searches that the Universe might actually be accelerating, have led many cosmologists to prefer a model with $\Omega_0 \approx 0.3$ and a cosmological constant that gives $\Omega_\Lambda = 0.7$ — a value chosen to make space flat (Box 2), thus satisfying advocates of inflation. This model appears to account better for measured microwave background fluctuations and large-scale structure, particularly new data on the evolution of galaxy clustering. So $\Omega_0 \approx 0.3$, $\Omega_\Lambda \approx 0.7$ has emerged, at least for the time being, as the 'new model Universe'. That is a remarkable occurrence, because many of us regard this model as exceptionally ugly, owing to the unexplained near coincidence of Ω_0 and Ω_Λ .

Could it nevertheless be the 'right' model? If it is, then for one thing it will have profound implications for particle physics. The existence of a cosmological constant (or vacuum energy) of the required size is not at all explained by current theories of the fundamental interactions of matter (see Box 2). Actually measuring Λ would provide an invaluable input for particle theorists to work on.

Although this appears to be the best-buy picture of the Universe at present, great uncertainty still clouds all three cosmological parameters. Nevertheless, there is every reason to be confident that the important issues will soon be resolved, because a data explosion is about to engulf cosmology. A new generation of huge galaxy surveys is already under way that can map the Universe out to depths where we should be able to find unambiguous clues to the nature and quantity of cosmological dark matter. The Anglo-Australian telescope has already begun to collect data for the 'Two-Degree Field', which will eventually map the positions of around a quarter of a million galaxies. The Sloan Digital Sky Survey will encompass more than a million galaxies, producing a vast database of galaxy properties. First light through the Sloan telescope has just been observed. And new satellite experiments, including the Microwave Anisotropy Probe (NASA) and the Planck Surveyor (ESA), will fly early in the next decade to scan the microwave background for the crucial Doppler peaks. The cosmological community is bracing itself for the arrival of these enormous new data sets and the new insight they will surely bring.

The question of whether our new model Universe will survive this data onslaught is less important than the realization that cosmology has at last come of age as an empiri-

cal science. Theoretical slack in the Big Bang model is being tightened all the time, and we should soon pin down the floating parameters with acceptable confidence. But perhaps none of the available family of models will fit all the new data. For many of us, that is the most exciting possibility of all, as we would have to move to stranger theories, perhaps not even based on General Relativity. □

Peter Coles is in the Astronomy Unit, Queen Mary and Westfield College, University of London, London E1 4NS, UK. He is Editor of The Icon Critical Dictionary of the New Cosmology (Icon Books, Cambridge, 1998).

e-mail: p.coles@qmw.ac.uk

1. Peebles, P. J. E., Schramm, D. N., Turner, E. L. & Kron, R. G. *Nature* **352**, 769–776 (1991).
2. Gott, J. R., Gunn, J. E., Schramm, D. N. & Tinsley, B. *Astrophys. J.* **194**, 543–553 (1974).
3. Blumenthal, G. R., Faber, S. M., Primack, J. R. & Rees, M. J. *Nature* **311**, 517–525 (1984).
4. Guth, A. H. *Phys. Rev. D* **23**, 347–356 (1981).
5. Guth, A. H. & Pi, S.-Y. *Phys. Rev. Lett.* **49**, 1110–1113 (1982).
6. Smoot, G. F. *et al. Astrophys. J.* **396**, L1–L4 (1992).
7. Freedman, W. L. *et al. Nature* **371**, 757–762 (1994).
8. Chaboyer, B., Demarque, P., Kernan, P. J. & Krauss, L. M. *Astrophys. J.* **494**, 96–110 (1998).
9. Coles, P. & Ellis, G. F. R. *Nature* **370**, 609–615 (1994).
10. Sandage, A. R. *Ann. Rev. Astron. Astrophys.* **26**, 561–630 (1988).
11. Garnavich, P. M. *et al. Astrophys. J.* **493**, L53–L56 (1998).
12. Perlmutter, S. *et al. Nature* **391**, 51–54 (1998).
13. Hancock, S., Rocha, G., Lasenby, A. N. & Gutiérrez, C. M. *Mon. Not. R. Astron. Soc.* **294**, L1–L6 (1998).
14. Caldwell, R. R., Dave, R. & Steinhardt, P. J. *Phys. Rev. Lett.* **80**, 1582–1585 (1998).
15. Weinberg, S. *Rev. Mod. Phys.* **6**, 1–22 (1989).

Hooke's housefly

The invention of the microscope opened up a new world of scientific discovery. It also presented perceptual and philosophical challenges – which were brought into sharp focus by the seminal work of Robert Hooke.

Martin Kemp

Every act of looking is an act of active interpretation. In our normal visual territories we unconsciously perform countless such acts each day without major problem. When we are confronted with unknown sights in visual landscapes of which we have no prior experience, the complex interaction between seeing and knowing becomes openly problematic. The arts of microscopic and telescopic observation present challenges to our perceptual abilities.

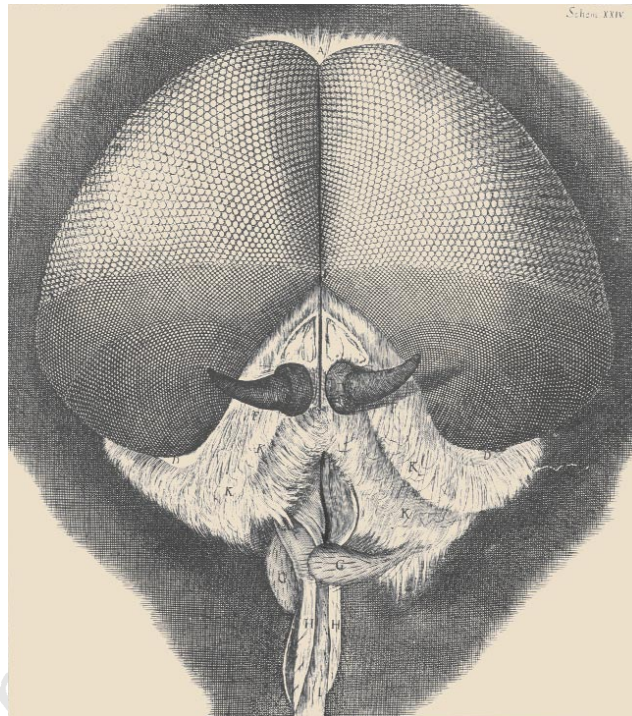
The perceptual issues and philosophical implications of microscopy are encapsulated in its first all-round masterpiece, Robert Hooke's *Micrographia*, published in London in 1665. Working in the context of the newly founded Royal Society, Hooke was commissioned to complete the project of microscopical observation and representation begun for King Charles II by Sir Christopher Wren, mathematician and architect.

Hooke was dedicated above all to "plainness and soundness of Observation". But, as he well knew, observing and representing are complex businesses, above all with the "adding of artificial Organs to the natural", as is the case with both microscopes and telescopes. The crucial problem is the "disproportion of the Object to the Organ" – whether the object is too big for the eye or too small, too close or too far away.

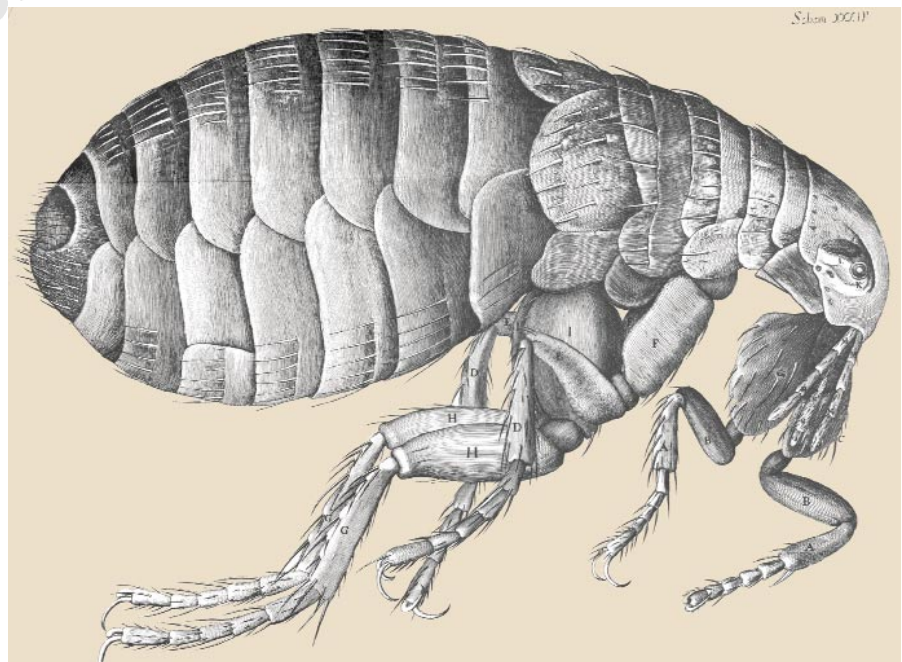
The key was to be able to translate the seen patterns of lights and darks into a coherent, three-dimensional image with reference to known forms. Hooke describes how "I have endeavoured.... first to discover the true appearance.... I never began to make a draught before by many examinations in several lights, and in several positions to these lights, I had discover'd the true form. For it is exceeding difficult in some Objects to distinguish between a prominency and a depression, between a shadow and a black stain, or a reflection and a whiteness in the colour."

As an example of the perceptual problems, he cites the eye of a fly, which was the subject of one of his most stunning plates: "The Eye of a Fly in one kind of light appears almost like a lattice, drill'd through with abundance of small holes... In the Sunshine they look like a surface cover'd with golden Nails; in another posture, like a surface cover'd with pyramids; in another with Cones; and in other postures of quite other shapes."

His text relies repeatedly on the use of



Hooke's "Eye of a Fly", from *Micrographia*, 1665.



Hooke's "Flea", from *Micrographia*, 1665.

analogies with the world of familiar objects. This use of resemblance serves to underline the ever more minute microcosmic affinities that microscopy was disclosing. As he wrote, "Little Objects are to be compar'd to the greater and more beautiful Works of Nature, A Flea, a Mite, a Gnat, to a Horse, an Elephant, or a Lyon". The flea, as a miracle of micro-engineering, is "adorn'd with a curiously polish'd suit of sable Armour, neatly

pointed, and beset with multitudes of sharp pins, shap'd almost like Porcupine's Quills, or bright conical Steel-bodkins".

Throughout the *Micrographia*, the beautiful mechanics and geometry of the smallest microcosms are made manifest, courtesy of Hooke's intelligent eye and elegant hand. □
Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.

Deadly relic of the Great War

The curator of a police museum in Trondheim, Norway, recently discovered in his archive collection a glass bottle containing two irregularly shaped sugar lumps. A small hole had been bored into each of these lumps and a glass capillary tube, sealed at its tip, was embedded in one of the lumps (Fig. 1). A note attached to the exhibit translated as follows: "A piece of sugar containing anthrax bacilli, found in the luggage of Baron Otto Karl von Rosen, when he was apprehended in Karasjok in January 1917, suspected of espionage and sabotage".

The curator noted the mention of anthrax with some alarm and he despatched the exhibit to the Norwegian Defence Microbiological Institute in Oslo which subsequently passed it to the Defence Evaluation Research Agency, Chemical and Biological Defence, Porton Down, UK.

Although 80 years had elapsed since the discovery of the sugar, the legendary durability of spores of *Bacillus anthracis* gave us reason to hope that some spores had survived, and we decided to try to revive them. To eliminate the chances of contamination of the sample, it was handled in a class 3 microbiological safety cabinet that had been fumigated with formaldehyde vapour and thoroughly cleaned.

We found no difficulty in removing the glass capillary tube from the sugar lump. The capillary was roughly 20 × 2 mm with tapered ends and contained a brown fluid. We opened it by scoring it with an ampoule file and cracking it over a sterile Petri dish. We transferred the liquid that remained in the capillary to a sterile 1.5 ml plastic tube, and estimated the volume of the liquid to be 10–15 µl. Liquid that spilled from the capillary was transferred to a separate sterile tube. The fragments of the capillary were also transferred to a tube and 1 ml of L-broth (tryptone 10 g, yeast extract 5 g, sodium chloride 5 g and water to 1 litre) was added. We also washed the Petri dish twice with 1 ml of L-broth and placed the washings in sterile tubes.

We incubated the washings and the glass fragments under static conditions at 37 °C for 8 days. After incubation, 200 µl of these cultures were spread on 7% horse-blood agar and L-agar medium (identical to L-broth but solidified by the addition of 2% Difco Bacto agar).

We cultured the liquid from the capillary directly onto agar because originally it must have contained a large number of spores to improve the chances of causing infection. We spread 2 × 0.5 µl and 1 × 1 µl volumes of the liquid onto 7% horse-blood agar and L-agar. If successful this would



Figure 1 Sugar lumps with an embedded glass capillary shown to contain spores of *Bacillus anthracis*.

have allowed us to determine the number of surviving spores. We incubated these plate cultures for 24 h at 37 °C, but none of them showed visible *B. anthracis* colonies.

We therefore resorted to enrichment, which is routinely used to revive apparently moribund bacteria. We added 2 µl of the original capillary suspension to 10 µl of L-broth and incubated the mixture at 37 °C for 8 days. We then spread 5 µl on 7% horse-blood agar to detect visible colonies. Eventually, therefore, all the liquid samples, washings and capillary fragments were subjected to the same enrichment procedure and culture on 7% horse-blood agar. We summarize the results in Table 1. Even after prolonged incubation in recovery media, the numbers of viable organisms that grew on subculture were extremely small, showing that the contents of the capillary were on the very edge of viability.

We carried out McFadyean's test on all the colonies to confirm that they were *B. anthracis*. In this test a subculture is made from the colony into a few millilitres of blood and the mixture is incubated for 4 h at 37 °C. A smear is then made on a microscope slide, fixed by heat or alcohol and stained with polychrome methylene blue. The rods of *B. anthracis* stain dark blue and are surrounded by a pink capsule; this confirmatory test provides incontrovertible identification of *B. anthracis*.

We also analysed the samples by polymerase chain reaction (PCR) as a back-up to the culture procedure. We placed 2 µl of the capillary contents into 10 µl of water in a 1.5 ml sterile Eppendorf tube and stored it overnight at 4 °C before analysis. We carried out separate PCR reactions on 2 µl aliquots of the resuspended sample, using primers specific for an anthrax-specific genomic

marker (*Ba*813), the lethal factor gene (*lef*) on plasmid pXO1 and the capsule B gene (*capB*) on plasmid pXO2. PCR reactions were prepared using Ready-To-Go PCR beads (Pharmacia Biotech) in a final volume of 25 µl. For each primer set, positive and negative control reactions were prepared using 10 ng of *B. anthracis* DNA or no DNA, respectively. A positive PCR containing *Ba*813 primers was spiked with 2 µl of sample to determine whether any PCR inhibitors were present.

The PCR thermal cycling conditions were: 95 °C for 5 minutes; 35 cycles at 95 °C for 30 seconds each; 55 °C for 30 seconds; and 72 °C for 60 seconds. The final extension time was increased to 10 minutes. We analysed PCR products by agarose gel electrophoresis at 80 volts for 1 hour on a 2% agarose gel containing 1 g ml⁻¹ of ethidium bromide. The gel was viewed over 302-nm ultraviolet light.

Spiking of a positive PCR with sample did not decrease amplification efficiency, showing that no PCR inhibitors were present. For all primer pairs, amplified DNA of the correct size was observed from the sample. This is consistent with the capillary containing pXO1⁺/pXO2⁺ *B. anthracis*.

We therefore confirmed the presence of *B. anthracis* in the specimen by both culture and PCR. It proved possible to revive a few surviving organisms from the brink of extinction after they had been stored, without any special precautions, for 80 years.

One of the most intriguing questions of this investigation is what Baron Otto Karl Robert von Rosen, a Swedish/German/Finnish aristocrat, and his companions were doing in this remote area of north-eastern Norway in the depths of winter; Karasjok is in Finnmark, about 18 kilometres from the Finnish frontier.

When they were arrested, all without passports, the contents of the baron's luggage were both revealing and compromising. When the Sheriff of Kautokeino, who was present at the group's arrest, derisively suggested that he should prepare soup from the contents of the tin cans labelled "Svea kott" (Swedish meat), the baron felt obliged to admit that each can actually contained between 2 and 4 kilograms of dynamite¹. Later, after the baron had been extradited from Norway, bottles of curare, "microbial cultures" and 19 sugar cubes, each contain-

Table 1 Recovery of colonies of *B. anthracis* after enrichment

Enriched sample	No. of colonies	McFadyean's test
Capillary contents	1	Positive
Capillary glass fragments	1	Positive
First wash	2	Positive
Second wash	0	N/A

ing a tiny glass tube holding anthrax microbes, were also identified. It seems, however, that the single specimen that we tested was the only one to have survived to the present day.

The baron claimed that he was an activist for Finnish independence and that their objective was to destroy lines of communication and transport, especially to Russian-controlled areas. According to his companions, the expeditions to these areas were in fact organized by the Germans and there had been several previous visits to Finland with sabotage equipment. Berlin had approved a proposal to use *B. anthracis* against reindeer used for sledging British arms through northern Norway².

The Norwegian newspapers at the time were convinced that the contaminated sugar cubes were to be used to disrupt the transport of merchandise by horses and reindeer between Skibotten and the Finnish border. The grinding of the sugar and its glass insert between the molar teeth of horses would probably result in a lethal infection as the anthrax spores entered the body, eventually facilitated through the small lesions produced in the wall of the alimentary tract by the broken glass.

It is not known whether reindeer eat sugar lumps but presumably the baron never had the chance to carry out this piece of research. Even if all the tubes had been used by the sabotage team it is unlikely that they would have caused much disruption to transportation, as anthrax is not transmissible directly from horse to horse. However, if the meat from a dying animal had been consumed without adequate cooking, it is likely that human fatalities from gastrointestinal anthrax would have followed.

At that time Norway's foreign policy was one of neutrality, but the commercial traffic across northern Norway was advantageous to the countries at war with Germany. This covert breach of neutrality, coupled with negotiations in progress in Kristiana (now Oslo) that were to bring Norway closer to the allies, resulted in the massive torpedoing of Norwegian ships in the Arctic Ocean.

After the baron was apprehended he was sent to Kristiana and held in custody for three weeks. He was then released and expelled to Sweden as a result of diplomatic pressure. Unfortunately the complete story was never officially discovered by the Norwegian authorities. It was clear that the baron was involved in acts of sabotage in Finland but the full nature of the contents of the baron's 'hunting equipment' was not revealed until after his expulsion. The disclosure of the dangerous biological agents in his luggage provoked heavy criticism of the Norwegian Ministry of Justice.

This small but relatively important episode in the history of biological warfare is one of the few instances of where there is

confirmation of the intent to use a lethal microorganism as a weapon, albeit 80 years after the event. It did not, however, make any significant difference to the course of the Great War.

Caroline Redmond, Martin J. Pearce, Richard J. Manchee

Detection Department, CBD, Porton Down, Salisbury SP4 0JQ, UK

Bjorn P. Berdal

Norwegian Defence Microbiology Institute, PO Box 4302 Torshov, N-0401 Oslo, Norway

1. Berg, R. *Norway on its Own, 1905–1920* (Univ. Oslo Press, 1995).
2. Hugh-Jones, M. Wickham Steed and German biological warfare research. *Intell. Nat. Security* 7, 379–402 (1992).

Cnidarian homeoboxes and the zootype

In a widely cited paper¹, Slack *et al.* proposed that all animal phyla shared a particular pattern of gene expression, the 'zootype'. The zootype hypothesis implies that at least six HOX-type homeobox-containing genes should be present in all metazoa. We have done a series of phylogenetic analyses which indicate that Cnidaria, thought to be the earliest-evolving animal phylum with the exception of the sponges², lack several HOX genes that are present in *Drosophila* and vertebrates. Instead, these genes may have arisen by duplication after the origin of the Cnidaria.

The zootype concept is based on the intriguing observations that, in members of diverse animal phyla, HOX genes are organized in clusters and the order of the genes within a cluster is related to the order of expression along the embryonic anterior–posterior axis. Slack *et al.*¹ proposed that the HOX gene expression pattern observed in triploblasts such as *Drosophila* and vertebrates 'is very ancient and was in place in the common ancestor of all modern animals'. To test this hypothesis,

we used a series of phylogenetic analyses to investigate the complement of HOX genes present in the Cnidaria, which is believed to be the earliest-originating animal phylum with an oral–aboral axis².

Screening for HOX genes in several cnidarian species has revealed the presence of a substantial number of putative HOX genes. To identify potential cnidarian HOX genes, we did an initial maximum parsimony analysis using a broad sampling of *Drosophila* and yeast homeobox genes³. We included all 24 cnidarian homeobox genes for which the complete homeobox region sequence is available. A second analysis focused on HOX genes, using the 14 cnidarian genes that form a monophyletic group with the *Drosophila* HOX genes; this analysis included *Drosophila* and mouse HOX genes and used sequence data for the conserved hexapeptide region present in many HOX genes, in addition to data for the homeobox region.

In the phylogenetic trees generated by both analyses, the genes in the middle of the HOX clusters form a monophyletic group that includes no cnidarian genes. This is most readily explained by derivation of these genes through duplication of a single precursor after the origin of Cnidaria. The anterior genes *labial* and *proboscipedia* and the posterior gene *Abdominal-B* branch earlier within the phylogenetic trees than the medially expressed genes. Thus, the trees are generally consistent with the presence in the common ancestor of Cnidaria and triploblasts of four HOX genes corresponding to one medial group precursor and the *Drosophila* genes *labial*, *proboscipedia* and *Abdominal-B* (Fig. 1).

In fact, four groups of cnidarian genes are present in the HOX trees, suggesting that four orthologous genes have been cloned from multiple cnidarian species. It is worth mentioning that the search for cnidarian homeobox genes has not been obviously biased against medial HOX genes. Most primers used in the polymerase chain reaction and library-screening probes

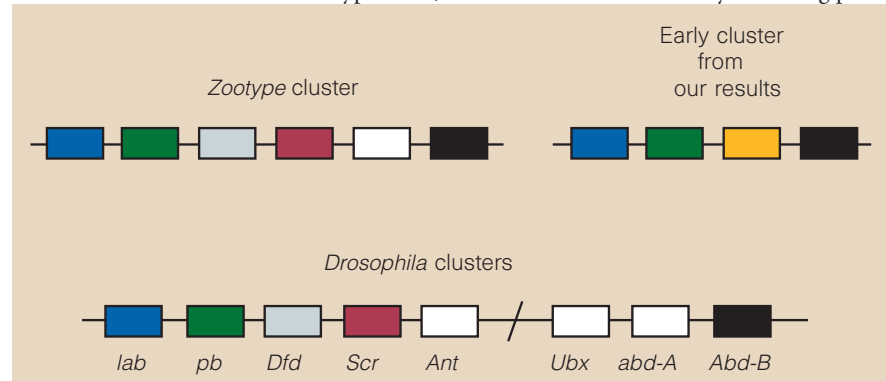


Figure 1 HOX gene cluster present in the common ancestor of all metazoans according to the zootype hypothesis, and HOX gene cluster present in the common ancestor of cnidarians and triploblasts according to our results, with *Drosophila* shown for comparison. Additional information on methods and results can be found at <http://pages.pomona.edu/~dmartinez/zootype.htm>

were designed to match the highly conserved region of helix 1 and/or helix 3 of the homeodomain of HOX genes⁴⁻⁸. Thus, the absence of certain types of HOX gene seems unlikely to be an artefact.

Previous phylogenetic results^{9,10} implied that, at an early stage in the evolution of the HOX cluster, there was a three-gene cluster consisting of an *Abdominal-B*-like gene, an ancestor of the medial genes, and an ancestor of the anterior genes *labial* and *proboscipedia*. Our results support a similar hypothesis, although they are consistent with the early presence of both anterior genes.

Our results do not support the zootype concept. Cnidarians appear to lack orthologues of several of the genes in the HOX clusters of *Drosophila* and vertebrates. Instead, these *Drosophila* and vertebrate genes form a monophyletic group, suggesting that they arose following the origination of Cnidaria. The zootype pattern of HOX gene expression is shared by a group within the metazoans, but it does not seem to provide "a morphological criterion for what an animal is"¹.

Daniel E. Martínez

Department of Biology, Pomona College,
Claremont, California 91711-6339, USA
e-mail: dmartinez@pomona.edu

Diane Bridge

Department of Biological Chemistry,
262 D Medical Sciences 1, University of California,
Irvine, California 92697-1700, USA

Liria M. Masuda-Nakagawa

Brain Science Institute, RIKEN, 2-1 Hirosawa,
Wako-shi, Saitama 351-0198, Japan

Pauly Cartwright

Harvard University Herbaria, Harvard University,
Cambridge, Massachusetts 02138-2094, USA

- Slack, J. M. W., Holland, P. W. H. & Graham, C. F. *Nature* **361**, 490-492 (1993).
- Schram, F. R. & Ellis, W. N. *Cladistics* **10**, 331-337 (1994).
- Swofford, D. L. PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.1 (Illinois Natural History Survey, Champaign, Illinois, 1993).
- Naito, M., Ishiguro, H., Fujisawa, T. & Kurosawa, Y. *FEBS Lett.* **333**, 271-274 (1993).
- Schummer, M., Scheurle, I., Schaller, C. & Galliot, B. *EMBO J.* **11**, 1815-1823 (1992).
- Shenk, M. A., Bode, H. R. & Steele, R. E. *Development* **117**, 657-667 (1993).
- Aerne, B. L., Baader, C. D. & Schmid, V. *Dev. Biol.* **169**, 547-556 (1995).
- Kuhn, K., Streit, B. & Schierwater, B. *Mol. Phy. Evol.* **6**, 30-38 (1996).
- Zhang, J. & Nei, M. *Genetics* **142**, 295-303 (1996).
- Agosti, D., Jacobs, D. & Desalle, R. *Cladistics* **12**, 65-82 (1996).

Making water levitate

The levitation in air of water, other diamagnetic substances¹ and even living organisms² was recently achieved by using the extremely strong magnetic field provided by a Bitter-type hybrid magnet. We too have

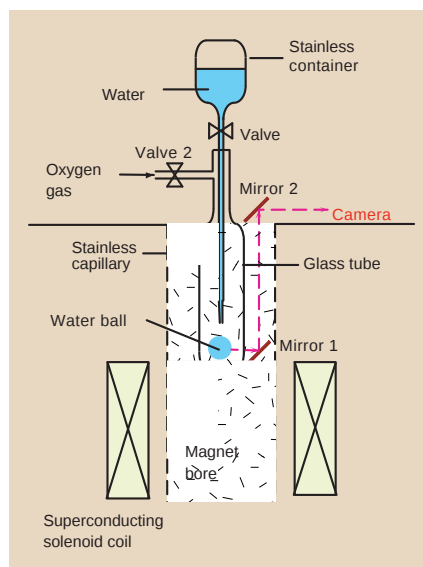


Figure 1 Set-up for the 'magneto-Archimedes levitation' of water in a pressurized oxygen atmosphere, $\rho_a = 1.21 \text{ kg m}^{-3}$, $\chi_a = +0.379 \times 10^{-6}$ for air and $\rho_w = 1.33 \text{ kg m}^{-3}$, $\chi_w = +1.80 \times 10^{-6}$ for oxygen; $\chi_w = -9.03 \times 10^{-6}$ at 20 °C and 1 atm, all in SI units.

succeeded in levitating water, but in the lower fields of an ordinary 10 T superconducting magnet. To achieve this we make use of gravitational and magnetically induced buoyancy forces in the host paramagnetic atmosphere (pressurized air or oxygen), rather than simply the diamagnetic force on the levitating object, to balance the gravitational force. This permits the magnetic levitation in air of paramagnetic as well as diamagnetic substances, which was widely believed to be impossible³. The physics underlying this effect is essentially the same as that of magnetohydrostatic ore separation, where a ferromagnetic fluid is used⁴. Because our process can levitate substances at a stable position in an atmosphere, we have named it 'magneto-Archimedes levitation'.

The presence of air inevitably contributes a small buoyancy force, due both to its weight (Archimedes' principle) and to the magnetic force that pulls the slightly paramagnetic air towards the centre of the magnet where the field is strongest. In the earlier experiments^{1,2}, the buoyancy forces on the levitating objects are likely to have

been small (~4% under the experimental conditions given) compared with the diamagnetic forces, and were not considered explicitly. But by increasing the pressure of the host atmosphere, these buoyancy forces can be amplified to the point where they dominate.

Including the Archimedes buoyant forces due to atmosphere of density ρ_a and magnetic volume susceptibility χ_a , the condition of stable levitation for water or a substance with ρ_w and χ_w is

$$-(\rho_w - \rho_a)g + [(\chi_w - \chi_a)/\mu_0]B \text{ dB/d}z = 0$$

where g is the acceleration due to gravity, μ_0 is the permeability of vacuum, B is the field intensity and z is the vertical position. The weight of air is still small relative to that of water, for instance at 60 atm. But more important is the ratio $|\chi_a/\chi_w|$, initially ~0.042 but then boosted to ~2.5 at 60 atm. Besides, χ is negative for water but positive for air, thereby adding up in the second term to a net enhancement of ~3.5. The magnet can then be weaker by this factor, reducing the field requirements to the range achievable by an ordinary superconducting magnet such as that used in the present experiment, of which the maximum $B \text{ dB/d}z = 420 \text{ T}^2 \text{ m}^{-1}$. As the paramagnetic susceptibility of the air is due to oxygen gas in the air, we could use pure oxygen gas to reduce the pressure required to 12 atm.

The demonstration was made with a set-up as shown in Fig. 1, using a cryo-cooler-operated 10 T magnet of 100 mm bore at room temperature (Sumitomo Heavy Industries; HF-10-100VHT). Distilled water, initially stored in the upper container, was injected through a stainless capillary into the glass tube vessel. After a water droplet had formed at the lower tip of the capillary, which was initially located at the height of the maximum $|B \text{ dB/d}z|$, the inserted container system was slowly retracted upwards as a whole so that the droplet was detached from the capillary.

For about the next minute the droplet repeated a complex bouncing motion, after which it stabilized to a free-standing spherical form. It was then possible to drip further water from above, one droplet after another, to enlarge the water ball to a vol-

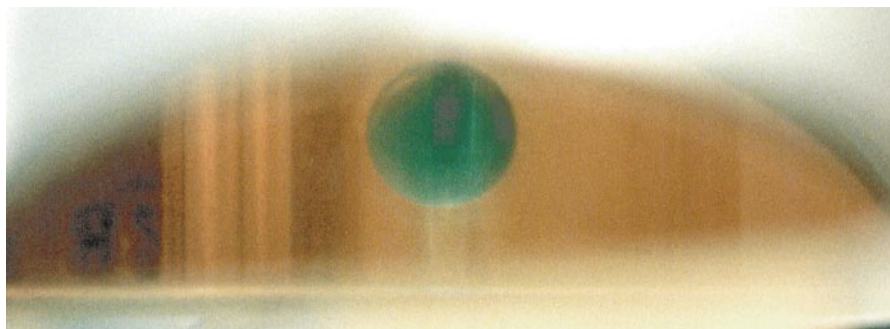


Figure 2 A levitating ball of aqueous CuSO_4 solution with $\chi_w = +0.48 \times 10^{-6}$ in SI units.

ume of $\sim 7 \text{ cm}^3$. Beyond this size the water ball was limited by the diameter of the glass container, touching its inner wall and spreading on the wall by wetting.

To demonstrate that the levitation is mainly due to the 'magneto-Archimedes' buoyant force, a paramagnetic water ball ($\chi = +0.48 \times 10^{-6}$ in SI units) containing CuSO_4 was levitated by the same procedures but under a little higher oxygen pressure, 18.4 atm. Figure 2 shows a photograph of a blue ball of paramagnetic aqueous CuSO_4 solution. To the best of our knowledge, there have been no reports of levitating a paramagnetic substance stably in the atmosphere. Furthermore, it has been claimed that it should be impossible to levitate a paramagnetic substance stably because a position cannot be created in the free space where the magnetic field is maximized according to one of the Maxwell relations, $\text{div} \mathbf{B} = 0$. But in magneto-Archimedes levitation the substance should float at the position of the minimum field horizontally, in the same position as in diamagnetic levitation.

Yasuhiro Ikezoe*, Noriyuki Hirota†, Jun Nakagawa‡, Koichi Kitazawa†

*Department of Applied Chemistry,
Department of Superconductivity,
University of Tokyo, 7-3-1 Hongo, Bunkyo-ku,
Tokyo 113-8656, Japan

†Japan Science and Technology Corporation,

4-1-8 Honcho, Kawaguchi, 332-0012 Japan

‡TDC Co., Ltd, Materials Research Center,
570-2 Matsugashita, Minamihatori, Narita,
286-8588 Japan

1. Beaunon, E. & Tournier, R. *Nature* **349**, 470 (1991).
2. Berry, M. V. & Geim, A. K. *Eur. J. Phys.* **18**, 307–313 (1997).
3. Jones, T. B. *J. Appl. Phys.* **50**, 5057–5058 (1979).
4. Khalafalla, S. E. & Reimers, G. W. Joint Meeting MMIJ-AIME, Tokyo, Print No. TIII e5 (1972).

Nicotine metabolism defect reduces smoking

Nicotine is the primary compound in tobacco that establishes and maintains tobacco dependence¹. Most of this nicotine is metabolized to cotinine by the genetically variable enzyme CYP2A6. Here we show that individuals lacking full functional CYP2A6, who therefore have impaired nicotine metabolism, are significantly protected against becoming tobacco-dependent smokers. In addition, smokers whose nicotine metabolism is thus impaired smoke significantly fewer cigarettes than those with normal nicotine metabolism.

Approximately one-third of the global population over 15 years of age smokes, despite the association of this activity with a higher incidence of many diseases including cancers and respiratory and cardiovascular

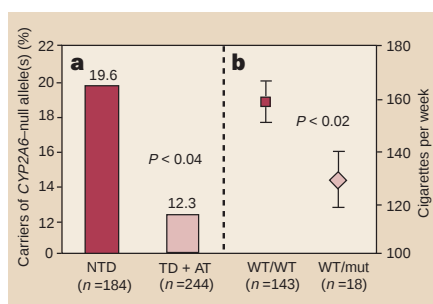


Figure 1 Effect of *CYP2A6*-null alleles on tobacco dependence. **a**, Percentage of individuals carrying *CYP2A6*-null alleles (1 or 2) in the never-tobacco-dependent exposure control group (NTD), and the combined tobacco-dependent (TD) and alcohol- and tobacco-dependent (AT) groups. **b**, Number of cigarettes smoked per week (mean $\pm$ s.e.m.) within the TD group is lower for individuals carrying a *CYP2A6*-null allele. No significant difference in age between groups ($P=0.05$); WT/WT: *CYP2A6**1/*1; WT/mut: *CYP2A6**1/*2 + *CYP2A6**1/*3.

disorders. In humans, 60–80% of nicotine that results from smoking is metabolized to cotinine², principally by the genetically polymorphic CYP2A6 enzyme^{3,4}. Three CYP2A6 alleles have been identified: the wild-type (*CYP2A6**1) and two null, or inactive, alleles (*CYP2A6**2 and *CYP2A6**3)⁵.

We considered that there would be an under-representation of individuals with impaired nicotine metabolism (carriers of null *CYP2A6* alleles) in a tobacco-dependent population. Genotype frequencies for CYP2A6 alleles⁵ were quantified in three groups: a tobacco-dependent group (TD, $n=164$, 52% male) defined by DSM-IV and Fagerstrom criteria, an alcohol- and tobacco-dependent group (AT, $n=80$, 84% male), and a never-tobacco-dependent exposure control group who had each tried smoking but had never become dependent (NTD, $n=184$, 52% male). Subjects in all groups had no other psychoactive drug dependencies (DSM-IV criteria). We found that, among all dependent smokers (TD + AT), the frequency of individuals with impaired nicotine metabolism (carriers of 1 or 2 *CYP2A6*-null alleles) was lower than in the control group (12.3% versus 19.6%, $P<0.04$, chi-square; odds ratio=1.74, 95% confidence intervals 1.02–2.94; Fig. 1a). These data show that impaired nicotine metabolism protects a smoker against becoming dependent: in fact, even a single *CYP2A6*-null allele (that is, heterozygosity) is sufficient to significantly reduce the risk of tobacco-dependence.

Dependent smokers adjust their smoking to maintain constant blood and brain nicotine concentrations^{6,7}. Therefore we tested whether dependent smokers with impaired nicotine metabolism would need to smoke fewer cigarettes. Within the TD group, those heterozygous for *CYP2A6*-null

alleles smoked significantly fewer cigarettes per week than smokers with two *CYP2A6* active alleles (129 versus 159 cigarettes per week; t -test $P<0.02$; Fig. 1b). These data suggest that CYP2A6-mediated nicotine metabolism is a significant determinant of the number of cigarettes consumed by dependent smokers and that heterozygosity in a single gene, the *CYP2A6* gene, significantly decreases this complex drug-taking behaviour.

Individuals carrying *CYP2A6*-null alleles may have a decreased risk of developing tobacco-related cancers⁸ and other medical complications because they have a decreased risk of becoming a smoker and, if they do become dependent, they smoke less than those without impaired nicotine metabolism. As tobacco smoke contains nitrosamines which can be activated to carcinogens by CYP2A6 (refs 9,10), individuals who carry *CYP2A6*-null alleles may also be less efficient at activating tobacco-smoke procarcinogens to carcinogens. These three factors may explain why there could be a reduction in tobacco related cancers for carriers of *CYP2A6*-null alleles.

CYP2A6 genotype may significantly affect nicotine levels from nicotine sources other than cigarettes, such as nicotine-replacement therapies (for example, patch, gum, spray). This may be of growing importance as such therapies are used increasingly for long-term maintenance against tobacco dependence and for treatment of other syndromes (such as Alzheimer's disease, Tourette's syndrome, ulcerative colitis). In addition, the protective effect of *CYP2A6*-null alleles against the risk of becoming tobacco-dependent and in decreasing consumption suggests that inhibiting this enzyme may be a new way to help prevent and treat tobacco smoking.

Michael L. Pianezza, Edward M. Sellers, Rachel F. Tyndale

Addiction Research Foundation of Ontario,
Centre for Research in Women's Health,
Department of Pharmacology, Medicine and
Psychiatry, University of Toronto, Toronto,
M5S 1A8, Canada
e-mail: r.tyndale@utoronto.ca

1. Henningfield, J. E., Miyasato, K. & Jasinski, D. R. *J. Pharmacol. Exp. Ther.* **234**, 1–12 (1985).
2. Benowitz, N. L. et al. *J. Pharmacol. Exp. Ther.* **268**, 296–303 (1994).
3. Nakajima, M. et al. *Drug Met. Disp.* **24**, 1212–1217 (1996).
4. Messina, E., Tyndale, R. F., & Sellers, E. M. *J. Pharmacol. Exp. Ther.* **282**, 1608–1614 (1997).
5. Fernandez-Salguero, P. et al. *Am. J. Hum. Genet.* **57**, 651–660 (1995).
6. McMorrow, M. J. & Fox, R. M. *Psychol. Bull.* **93**, 302–327 (1983).
7. Russel, M. S. H. in *Advances in Behavioural Biology* (eds Martin, W. R., Vauhan, G. R., Iwamoto, E. T., & David, L.) 25–50 (Plenum, New York, 1987).
8. Law, M. R., Morris, J. K., Watt, H. C., & Wald, N. J. *Brit. J. Cancer* **75**, 1690–1693 (1997).
9. Crespi, C. L. et al. *Carcinogenesis* **11**, 1293–1300 (1990).
10. Yamazaki, H. et al. *Carcinogenesis* **13**, 1789–1794 (1992).

History under the microscope

Instruments of Science: An Historical Encyclopedia

edited by Robert Bud and Deborah Jean Warner

Garland: 1998. Pp. 709. £100, \$187.50

Arthur Middleton

When I was asked to review this book, I accepted happily enough. Later, I had misgivings: a comprehensive encyclopedia of scientific instruments? Most encyclopedias, unless they are the size and scope of the *Encyclopaedia Britannica*, miss out a good deal. My first queries were: "does it start from the first instruments that were used by man?", and then "is it up to date, including the instruments used in early experiments into nuclear physics?" I am relieved to report that it does all this, and more. Up-to-date entries include the American rail track loading vehicle of 1990, the ion-sensitive microelectrode from the 1970s, and the laser and maser.

The idea for this book germinated five years ago, leading to a collaboration between the Science Museum, London, and the Smithsonian Institution, Washington DC. The 223 contributors who wrote the 327 entries volunteered their time and expertise, or were otherwise pushed, cajoled, bullied, entreated or pleaded to contribute, and so they did, but not for money, receiving a free copy of the encyclopedia instead. The entries in many cases are subdivided to take into account the development of a particular instrument over decades, such as the chronoscope, or over centuries, such as the telescope, from the earliest example of 1608 to the modern radio- and X-ray telescopes.

The illustrations are mostly photographs, with a scattering of prints from early books (from the fifteenth to the eighteenth century) or makers' catalogues from the early nineteenth century onwards. Other illustrations are schematic. Where the photograph of a larger instrument includes the operator standing beside it, some scowling, as in the old days of photography (like Michel Ter-Pogossian with an early positron emission tomography scanner), or grinning broadly (like John Mallard with his prototype magnetic resonance imaging device), there is a good idea of size and scale. But I wish that the other photographs, particularly those of laboratory instruments, included their measurements in the caption. Jesse Ramsden's three-foot geodetic theodolite is a large and splendid item, but William Abney's level would sit comfortably in the palm of the hand: one is left to guess.

The encyclopedia's advisory committee of nine includes such impressive names as Robert Anderson of the British Museum, Paolo Brenni, recently elected vice-chairman

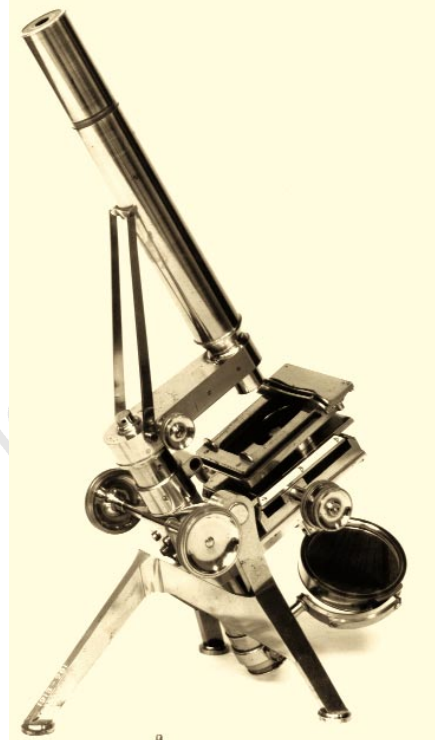


Instrumental to science history: a planispheric astrolabe from 1666 (above) and a Powell and Lealand achromatic telescope from 1846 (right).

of the Scientific Instruments Commission of the International Union of the History and Philosophy of Science, and Gerard Turner, former president of the same and president of the Scientific Instrument Society. All but three of the committee also contributed.

Of the 217 others, I was particularly pleased to see that Diana Crawforth-Hitchens contributed a piece about the general balance, as she has become an authority on scales and balances since the untimely death of her husband a few years ago. Jim Bennett, of the Museum of the History of Science at the University of Oxford, writes clearly on early surveying and navigation instruments, subjects covered in his book *The Divided Circle* (now, alas, out of print). David King of Frankfurt reveals the workings of the planispheric astrolabe in an understandable manner, although the mariner's version is illustrated by a seventeenth-century print; a photograph of an original would have been better. There are only 80 authenticated examples in the world, but some of them, despite immersion in the sea for hundreds of years, still have remarkably fresh divisions and numbering. They survived so well because they were wrapped in cloth and kept in wooden chests, before being brought up from the deep along with the treasure. Others were dropped overboard, probably by the numb, cold hands of a profane seaman, and much later a scoured fragment would turn up in a fisherman's net.

This book will be of greatest value to students of the history of science. How many



young people know the difference between the mariner's astrolabe and the planispheric (or Persian) type, what they look like and how they work; can tell at a glance if a telescope is a reflector or a refractor; or why the sextant and the octant are so called, when they both do the same job in the same manner? Now they might find out.

Turning to the first entry, the abacus, I was encouraged to find a lucid explanation of the differences between the Western and Eastern styles of the instrument, something difficult to find elsewhere. The book also explains that the royal accounts of King Henry II of England (1154–89) were calculated over a table called a chessboard, hence the expression 'the exchequer'.

Comparing the numbers of instruments to the time frame in which they were invented, it is not surprising that the nineteenth and twentieth centuries have most, 115 each. This then declines rapidly as we go back in time, to 37 in the eighteenth century, 30 in the seventeenth century and 31 before that. The entries are between 800 and 1,800 words, hence the need for 700 pages.

The short bibliography at the end of each subject is useful, and in many cases one of the reference works cited was written by the contributor, such as that about the patch-clamp amplifier by Erwin Neher, the sphygmomanometer (which measures arterial blood pressure) by Hughes Evans, or the equatorium (a Renaissance calculator that provides angular coordinates to determine the true longitude of a planet) by Mercè Comes. The

equatorium is a rare and complex instrument, and is not easy for non-astronomers to understand. The example illustrated, one of only a handful known, is in the collection of the National Museums and Galleries at Merseyside, in the United Kingdom. A cropped version is used as the colour illustration on the front cover, but very few readers would recognize it. A better, and perhaps more symbolic, choice might have been the Powell and Lealand microscope of 1846. Just 60 years later, Charles Rolls and Henry Royce started making cars, in the same limited quantity but to the same exacting standards.

One of my sons is presently reading for his MSc. I considered giving him my review copy of this encyclopedia but, on second thoughts, I shall make good use of it myself. □

Arthur Middleton, a dealer in the history of science, is at 12 New Row, Covent Garden, London WC2N 4LF, UK.

Darwin's fixed course

Evolution: Society, Science and the Universe

edited by A. C. Fabian
Cambridge University Press: 1998. Pp. 179.
£16.95, \$22.95

Mark Pagel

The irony of the modern tendency to use the word 'evolution' as synonymous with natural selection is that, like the traits of many species, it may reveal something of its ancestry. In Darwin's time, 'evolve' meant to unfold, roll out or unfurl, suggesting a fixed course. He anxiously wished to avoid this meaning for his new-found 'natural selection', and so it may not be surprising that he used 'evolve' only once in the first edition of *The Origin of Species*, and even then he waited until the very end, to the last word of the book. And yet, some of the products of natural selection are so exquisite as to suggest something preordained, and 'evolve' in its old denotative sense finds its way back into our discourse.

Meanings change, however. The contemporary scientific concept of evolution as simply 'to change over time' is ever more widely applied, and not just to refer to living things. A. C. Fabian's delightful new edited volume is a case in point. Fabian invited eight well-known "communicators from separate disciplines" to discuss evolution in the successful Darwin College lecture series at the University of Cambridge. The result is an eclectic collection that can fail to interest only the modern philistine.

The possibility that forms of natural selection lurk as 'invisible hands' in areas well removed from biology does not always settle easily in the mind of the essayist. In her

Face value

The arrangement of features in the human face largely reflects sensory, dietary and linguistic considerations, so all faces look more or less the same. This means that minor variations can effectively convey social signals and information about identity. This is why natural facial disfigurements have such a profound effect on people's lives. It also explains the disturbing effect of artistic violations of the face, such as this striking painting, *Le Viol (The Rape)* by René Magritte. Or so argue Vicki Bruce and Andy Young in *In the Eye of the Beholder: The Science of Face Perception* (Oxford University Press, £25). They bring together science and art to explain the importance of the face, how we extract the information it contains, and reveal what it all means.



chapter, Gillian Beer allows that readers and novels have evolved. She is less willing to countenance the view that works outside the literary canon, although less 'fit', are somehow less worth reading. The architect Richard Rogers argues passionately that London has been in steady decline since the mid-1980s when the government removed its central planning authority. Rogers and Beer are making a point well known to biologists: that which emerges from the unsentimental struggle for survival is the fittest, but not necessarily the most desirable.

Biologists wishing to understand human civilization from a darwinian perspective confront a dilemma, avers Tim Ingold. All human cultural evolution has taken place against the same biological background: Fred Flintstone, had he been born later, might have been Shakespeare or Einstein. This, Ingold says, forces biologists to adopt a 'progressive' view of history as the "unfolding of pre-evolved potentials", with 'primitive' hunter-gatherers at the bottom and Western scientists at the top: a "fundamentally teleological view!" Ingold exclaims triumphantly.

But cultural evolution is not a one-way street: cultures adopt more 'primitive' subsistence practices when it is advantageous to do so and, as Jared Diamond's informative chapter summarizes, they lose as well as acquire technologies. He records how Tasmanian islanders lost the ability to fish and even to start fires following their geographical separation from mainland Australians.

Ingold's position shows how difficult it can be to translate the arguments and assumptions of one field to another.

Still, evolutionary processes may unfold in more subtle ways than is realized. The astronomer Martin Rees reluctantly ponders whether our Universe could have been much different from what we see, given the laws of physics that emerged with the Big Bang. A similar question worries biologists: are the forms we see now (such as kangaroos, toadstools, fruitflies and sequoias) mere accidents, or if we were to re-run evolution would similar (if not the same) kinds of organisms emerge? Intriguingly, evidence is pointing towards the latter view, confounding many a sceptic. And what of cities and even novels? The future may not be preordained but its broad outlines might be.

The long march of evolution now includes spin-offs into other fields; E. O. Wilson sees this as a process of 'conciliance'. This was not lost on Fabian, who, having brought his speakers to Cambridge, produced this fine volume. With contributors ranging from those discussed here to Freeman Dyson, Lewis Wolpert and Stephen Jay Gould, Fabian's book reaffirms my view that the best researchers and scientists often make the best science writers: no exaggerated claims, no mysterious theories of everything, no yin and yang. Just a joy to read. □

Mark Pagel is in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

Two feathered dinosaurs from northeastern China

Ji Qiang*, Philip J. Currie†, Mark A. Norell‡ & Ji Shu-An*

* National Geological Museum of China, Yangrou Hutong 15, Xisi, 100034 Beijing, People's Republic of China

† Royal Tyrrell Museum of Palaeontology, Box 7500, Drumheller, Alberta T0J 0Y0, Canada

‡ American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024-5192, USA

Current controversy over the origin and early evolution of birds centres on whether or not they are derived from coelurosaurian theropod dinosaurs. Here we describe two theropods from the Upper Jurassic/Lower Cretaceous Chaomidianzi Formation of Liaoning province, China. Although both theropods have feathers, it is likely that neither was able to fly. Phylogenetic analysis indicates that they are both more primitive than the earliest known avialan (bird), *Archaeopteryx*. These new fossils represent stages in the evolution of birds from feathered, ground-living, bipedal dinosaurs.

Dinosauria Owen 1842
Theropoda Marsh 1881
Maniraptora Gauthier 1986
Unnamed clade
Protarchaeopteryx robusta Ji & Ji 1997

Holotype. National Geological Museum of China, NGMC 2125 (Figs 1, 2 and 3).

Locality and horizon. Sihetun area near Beipiao City, Liaoning, China. Jiulongssong Member of Chaomidianzi Formation, Jehol Group¹. This underlies the Yixian Formation, the age of which has been determined to be Late Jurassic to Early Cretaceous^{3,4}.

Diagnosis. Large straight premaxillary teeth, and short, bulbous maxillary and dentary teeth, all of which are primitively serrated. Rectrices form a fan at the end of the tail.

Description. The skull of *Protarchaeopteryx* is shorter than the femur (Table 1). There are four serrated premaxillary teeth (Fig. 1c), with crown heights of up to 12 mm. Premaxillary teeth of coelophysids⁵, compsognathids^{6,7} and early birds lack serrations, but premaxillary denticles are present in most other theropods. Six maxillary and seven dentary teeth are preserved (Fig. 1), all of which are less than a quarter the height of the premaxillary teeth. They most closely resemble those of *Archaeopteryx*⁸ in shape (Figs 1b, c and 2b, c), but have anterior and posterior serrations (7–10 serrations per mm).

The amphicoelous posterior cervicals are the same length as the posterior dorsals, which have large pleurocoels. If the lengths of missing segments of the tail are accounted for, there were fewer than 28 caudals. Vertebrae increase in length from proximal to mid-caudals, as in most non-avian coelurosaurs.

There are two thin, flat, featureless sternal plates. The clavicles are fused into a broad, U-shaped furcula (interclavicular angle is about 60°) as in *Archaeopteryx*, *Confuciusornis* and many non-avian theropods. The forelimb is shorter than the hindlimb. The arm is shorter (compared to the femur) than it is in birds, but is longer than those of long-armed non-avian coelurosaurs such as dromaeosaurids and oviraptorids (Table 2). The better preserved right wrist of NGMC 2125 has a single semilunate carpal capping the first two metacarpals. The hand has the normal theropod phalangeal formula of 2-3-4-x-x. The manus is longer than either the humerus or radius. Compared to femur length, the hand is more elongate than those of any theropods other than *Archaeopteryx*⁹ and *Confuciusornis* (Table 2). More advanced birds such as *Cathayornis* have shorter hands¹⁰. Phalanges III-1 and III-2 in the hand of *Protarchaeopteryx* are almost the same size, and are about half the length of III-3. The unguals are long and sharp, and keratinous sheaths are preserved on two of them.

The preacetabular blade of the ilium is about the same length as the postacetabular blade. The pubic boot expands posteriorly. Anteriorly, the pubis is not exposed.

The tibia is longer than the femur, as it is in most advanced theropods and early birds. It is not known if the fibula extended to the tarsus.

The metatarsals are separate from each other and the distal tarsals. Metatarsal I is centred halfway up the posteromedial edge of the second metatarsal. In perching birds such as *Sinornis*³, metatarsal I is positioned near the end of metatarsal II and is retroverted. Its condition in *Archaeopteryx* is intermediate. Pedal unguals are smaller than manual unguals.

A clump of at least six plumulaceous feathers is preserved anterior

Table 1 Lengths of elements in *Protarchaeopteryx* and *Caudipteryx*

Element	NGMC 2125	NGMC 97-4-A	NGMC 97-9-A
Body length	690	890	725
Skull	70	76	79
Sternal plates	25	36	—
Humerus	88	69	70
Arm (humerus to end of phalange II-2)	297	214	220
Ilium	95	101	—
Ischium	—	77	—
Leg (femur to end of phalange III-4)	450	550	540
Femur	122	147	149
Tibia	160	188	182
Tibia	160	188	182
Metatarsal III	85	115	117

Length measurements are given in millimetres. NGMC 2125, *Protarchaeopteryx*; NGMC 97-4-A and NGMC 97-9-A, *Caudipteryx*.

Table 2 Relative proportions of elements in relevant avian and non-avian theropods

Element	Drom	Ov	Tro	Cx	Px	Ax	Con
Arm/F	1.8–2.6	1.5–1.8	1.8	1.5	2.4	3.7	3.9
S/H	0.8	1.0–1.2	—	1.1	—	0.6	0.8
R/H	0.7–0.8	0.8–0.9	0.6–0.7	0.9	0.8	0.9	0.8
Manus/H	0.9–1.2	1.2–1.4	1.3	1.2	1.6	1.2	1.3
Manus/F	1.0	0.7–1.0	0.8	0.6	1.2	1.5	1.6
Mcl/McII	0.4–0.5	0.4–0.6	0.3	0.4	0.4	0.3	0.4
T/F	1.1–1.4	1.2	1.1–1.2	1.2	1.3	1.4	1.1
Leg/F	3.6	3.3	3.8	3.7	3.7	3.8	3.3
Leg/arm	1.4	1.7	2.1	2.5	1.5	1.1	0.8

All data were collected from original specimens by P.J.C. Ax, *Archaeopteryx*; Con, *Confuciusornis*; Cx, *Caudipteryx*; Drom, dromaeosaurids; F, femur; H, humerus; L, length; Mc, metacarpal; Ov, oviraptorids; Px, *Protarchaeopteryx*; R, radius; S, scapula; T, tibia; Tro, troodontids.

to the chest, with some showing well-developed vanes (Fig. 3a). Evenly distributed plumulaceous feathers up to 27 mm long are associated with ten proximal caudal vertebrae. Twenty-millimetre plumulaceous feathers are preserved along the lateral side of the right femur and the proximal end of the left femur.

Parts of more than twelve rectrices are preserved¹¹ attached to the distal caudals. One of the symmetrical tail feathers (Fig. 3b) extends 132 mm from the closest tail vertebra, and has a long tapering rachis with a basal diameter of 1.5 mm. The well-formed pennaceous vanes of *Protarchaeopteryx* show that barbules were present. The vane is 5.3 mm wide on either side of the rachis. At midshaft, five barbs come off the rachis every 5 mm (compared with six in *Archaeopteryx*), and individual barbs are 15 mm long. As in modern rectrices, the barbs at the base of the feather are plumulaceous.

Maniraptora Gauthier 1986
Unnamed clade

Diagnosis. The derived presence of a short tail (less than 23 caudal vertebrae) and arms with remiges attached to the second digit.

Caudipteryx zoui gen. et sp. nov.

Etymology. ‘*Caudipteryx*’ means ‘tail feather’, ‘*zoui*’: refers to Zou

Jiahua, vice-premier of China and an avid supporter of the scientific work in Liaoning.

Holotype. NGMC 97-4-A (Figs 4 and 5b).

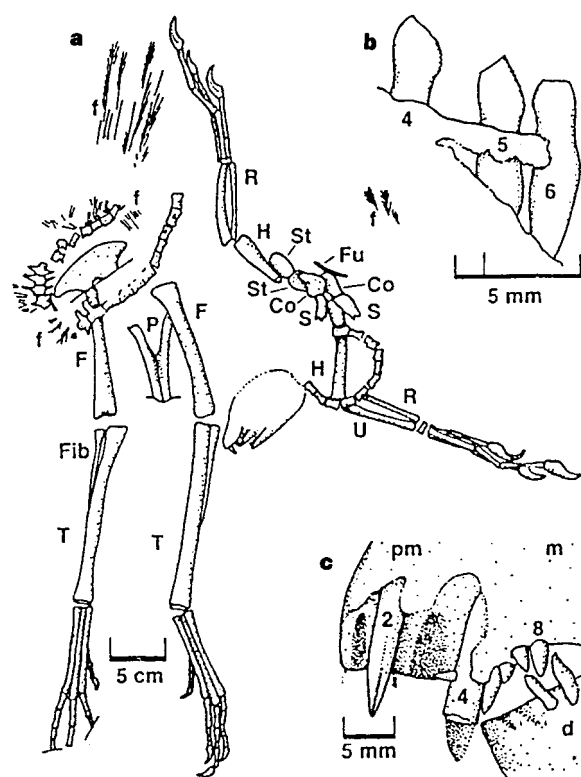
Paratype. NGMC 97-9-A (Fig. 5d).

Locality and horizon. Sihetun area, Liaoning. Jiulongsong Member of the Chaomidianzi Formation.

Diagnosis. Elongate, hooked premaxillary teeth with broad roots; maxilla and dentary edentulous. Tail short (one-quarter of the length of the body). Arm is long for a non-avian theropod; short manual claws. Leg-to-arm ratio, 2.5.

Description. The skulls of both specimens of *Caudipteryx* are shorter than the corresponding femora because of a reduction in the length of the antorbital region. The relatively large premaxilla (Figs 6 and 7) borders most of the large external naris. The maxilla and nasal are short, but the frontals and jugals are long. The lacrimal of NGMC 97-4-A is an inverted L-shaped, pneumatic bone. Scleral plates are preserved in the 20-mm-diameter orbits of both specimens. The tall quadratojugal seems to have contacted the squamosal and abutted the lateral surface of the quadrate. The single-headed quadrate is vertical in orientation. The ectopterygoid has a normal theropod hooklike jugal process. There is a broad, beak-like margin at the symphysis of the dentaries. Posteriorly, the dentary bifurcates around a large external mandibular fenestra as in oviraptorids. A





◀ **Figure 2** *Protarchaeopteryx robusta*. **a**, Outline of the specimen shown in Fig. 1a. **b**, Outline of the left dentary teeth shown in Fig. 1b. **c**, Drawing of the front of the jaws, showing the large size of the premaxillary teeth compared with maxillary and dentary ones. Abbreviations: Co, coracoid; d, dentary; F, femur; f, feathers; Fib, fibula; Fu, furcula; H, humerus; m, maxilla; P, pubis; pm, premaxilla; R, radius; S, scapula; St, sternal plate; T, tibia; U, ulna. Numbers represent tooth positions from front to back.



Figure 1 *Protarchaeopteryx robusta*. **a** (facing page), NGMC 2125, holotype. Scale bar, 5 cm. **b**, Fourth to sixth left dentary teeth. Scale bar, 1 mm. **c**, Premaxillary teeth showing small serrations. Scale bar, 5 mm.

well-developed, sliding intramandibular joint is present between dentary and surangular.

There are four teeth in each premaxilla. They have elongate, needlelike crowns, and the roots are five times wider than the crowns (Fig. 7b). The lingual wall of the root of the third right tooth has been resorbed for the crown of a replacement tooth. The teeth seem to have been procumbent, with an inflection at the gumline. *Caudipteryx* had no maxillary or dentary teeth.

There are ten amphicoelous cervical vertebrae and five sacra as in most non-avian theropods and *Archaeopteryx*^{8,12}. The tail of NGMC 97-4-A is articulated and well-preserved, and includes 22 vertebrae, as in *Archaeopteryx*. It is shorter than the 30-segment tails of oviraptorids. Most other non-avian theropods have much longer tails. Caudals do not become longer posteriorly, as they do in most non-avian theropods and *Archaeopteryx*. Almost two-thirds of the tail of NGMC 97-4-A is preserved as a straight rod, but the vertebrae are not fused. The first six haemal spines are elongate, rodlike structures. More posterior haemal spines decrease in height, but expand anteriorly and posteriorly (Fig. 5a).



Each segment of gastralia is formed by two pairs of slender, tapering rods, as in all non-avian theropods, *Protarchaeopteryx* and early birds^{8,9,13,14}.

The paired sternals are similar to those of dromaeosaurids and oviraptorids. *Confuciusornis* had a relatively larger, unkeeled sternum. Some short bones with slight expansions at each end are found near the sternal plates of NGMC 97-9-A, and may be sternal ribs.

The scapula is longer than the humerus, whereas the scapula-humerus ratio is less than 1.0 in flying birds (Table 2) because of humerus elongation. The clavicles are fused into a broad, U-shaped furcula in NGMC 97-9-A as in *Archaeopteryx*, *Confuciusornis* and many non-avian theropods.

Compared to the humerus, forearm length is similar to that in oviraptorosaurs (Table 2), *Archaeopteryx* and *Protarchaeopteryx*. In more advanced birds^{15,16}, the radius is longer than the humerus. The external surface of the ulna, as in *Archaeopteryx*⁸, lacks any evidence of quill nodes.

There are three carpals preserved in NGMC 97-4-A, including a large semi-lunate one that caps metacarpals I and II as in dromaeosaurids, oviraptorids, troodontids, *Archaeopteryx*, *Confuciusornis* and other birds. Four carpals have been recognized in *Archaeopteryx*¹². A large triangular radiale sits between the semi-lunate and the radius. A small carpal articulates with the third metacarpal. A thin wedge of bone at the end of the ulna is probably a fragment of gastralia.

The unfused metacarpals and digits of both specimens are well preserved. The third metacarpal is almost as long as the second, but is more slender. The hand has the normal theropod phalangeal



Figure 3 *Protarchaeopteryx robusta*, NGMC 2125. **a**, Contour and plumulaceous feathers. Scale bar, 10 mm. **b**, Rectrices. Scale bar, 5 mm.

formula of 2-3-4-x-x. The manus is longer than either the humerus or the radius, which is a primitive characteristic shared with most non-avian coelurosaurs, *Archaeopteryx*⁹ and *Confuciusornis*. In contrast with *Archaeopteryx*, *Confuciusornis*, *Protarchaeopteryx* and many non-avian theropods (ornithomimids, troodontids, dromaeosaurids and oviraptorids), the manus is relatively short compared with the femur.

The curved second manual ungual is about two-thirds the size of the same element in *Protarchaeopteryx*, and is less than 70% the length of the penultimate phalanx.

Pelvic elements are unfused, as they are in all non-avian theropods (except some ceratosaurs) and the most primitive birds¹⁶. The acetabulum is large, comprising almost a quarter of the length of the ilium (the ratio of acetabulum-to-ilium length is less than 0.11 in birds¹⁷). It has a deeper, shorter, more squared-off pre-

acetabular region than that of *Protarchaeopteryx*, and closely resembles the ilium of dromaeosaurids¹⁸. The tapering post-acetabular region is lower and longer than the preacetabular. The pubic penduncle is anteroposteriorly elongated, and has a notch (Figs 4 and 5b) in the ventral margin that divides the suture into two surfaces. This notch and the deep pubic penduncle of the ischium are characteristic of opisthopubic pelvises. The ischium has no dorsal process such as that found in *Archaeopteryx* and *Confuciusornis*, and the shaft curves down and back. A well-developed ventromedial flange is present, perhaps indicating contact between elements. In general appearance, the ischium most closely resembles those of non-avian coelurosaurs.

The ratio of hindlimb-to-forelimb length is higher than in other coelurosaurs (Table 2) except alvarezsaurids¹⁹, which had exceptionally short arms. The greater trochanter is separated from the



Figure 6 *Caudipteryx zoui*, skull of NGMC 97-9-A in right lateral view. Scale bar, 1 cm.

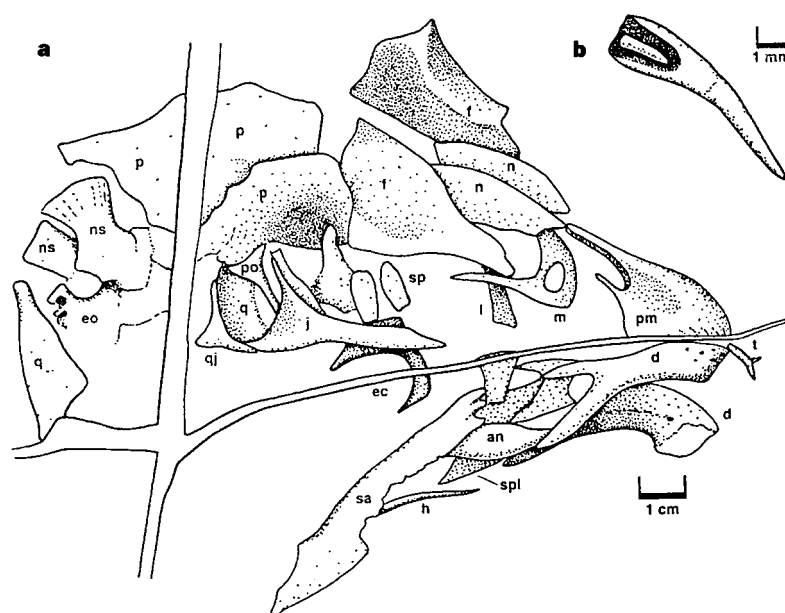


Figure 7 *Caudipteryx zoui*. **a**, Sketch of skull shown in Fig. 6. **b**, Premaxillary tooth of NGMC 97-4-A, showing resorption pit and germ tooth. Abbreviations: an, angular; d, dentary; ec, ectopterygoid; eo, exoccipital; f, frontal; h, hyoid; j, jugal; l,

lacrymal; m, maxilla; n, nasal; ns, neural spine; p, parietal; pm, premaxilla; po, postorbital; q, quadrate; qj, quadratojugal; sa, surangular; sp, scleral plate; spl, splenial; t, premaxillary teeth.

lesser trochanter of the femur by a shallow notch, and forms a raised, semi-lunate rim that is similar to the trochanter femoris of birds, troodontids and avimimids.

None of the fibulae is complete, but NGMC 97-9-A has a socket for the distal end of the fibula formed by the calcaneum, astragalus and tibia. The astragalus is not fused to the tibia. The ascending process of NGMC 97-9-A (Fig. 5e) extends 22% of the distance up the front surface of the tibia, compared with 12% in *Archaeopteryx*¹². As in *Archaeopteryx*⁸, *Confuciusornis*¹⁰ and most non-avian theropods, the calcaneum is retained as a separate, disk-like element. Two distal tarsals are positioned over the third and fourth metatarsals, as in *Archaeopteryx*, *Buluochia*²⁰ and all non-avian theropods that lack fused tarsometatarsals.

The metatarsals of *Caudipteryx* are not fused; this is the plesiomorphic condition expressed in most non-avian theropods. Metatarsal I is centred about a quarter of the way up the posteromedial corner of the second metatarsal. The third is the longest of the metatarsals, and in anterior view completely separates the second and fourth metatarsals, unlike in the arctometatarsalian condition of many theropods²¹. Nevertheless, at midshaft the third metatarsal is thin anteroposteriorly and is triangular in cross-section. The pedal unguals are triangular in cross-section and are about the same size as the manual unguals.

At least fourteen remiges are attached to the second metacarpal,

phalanx II-1, and the base of phalanx II-2 of NGMC 97-4-A (Fig. 8a). Each remex has a well-preserved rachis and vane. The most distal remex is less than 30 mm long. The second most distal remex is 63.5 mm long, is symmetrical, and has 6.5-mm-long barbs on either side of the rachis. The fourth most distal primary remex is 95 mm long and is longer than the humerus. Unfortunately, the distal ends of the remaining remiges are not preserved. In flying birds (even *Archaeopteryx*¹²), each remex is longer than they are in *Caudipteryx*, and the most distal remiges are the longest. For example, the remiges of *Archaeopteryx*²² are more than double the length of the femur. The barbs on either side of the rachis are symmetrical, contrasting with *Archaeopteryx* and modern flying birds²³.

The holotype preserves ten complete and two partial rectrices. Eleven are attached to the left side of the tail and were probably paired with another eleven feathers on the right side (only the terminal feather is preserved). Two rectrices are attached to each side of the last five or six caudal vertebrae, but not to more anterior ones. NGMC 97-9-A preserves most of nine rectrices. In *Archaeopteryx*, rectrices are associated with all but the first five or six caudals^{12,22}. Each rachis has a basal diameter of 0.74 mm and tapers distally. All the feathers appear to be symmetrical (Fig. 8b), although in most cases the tips of the barbs of adjacent feathers overlap. The vane of the sixth feather is 6 mm wide on either side of the rachis.

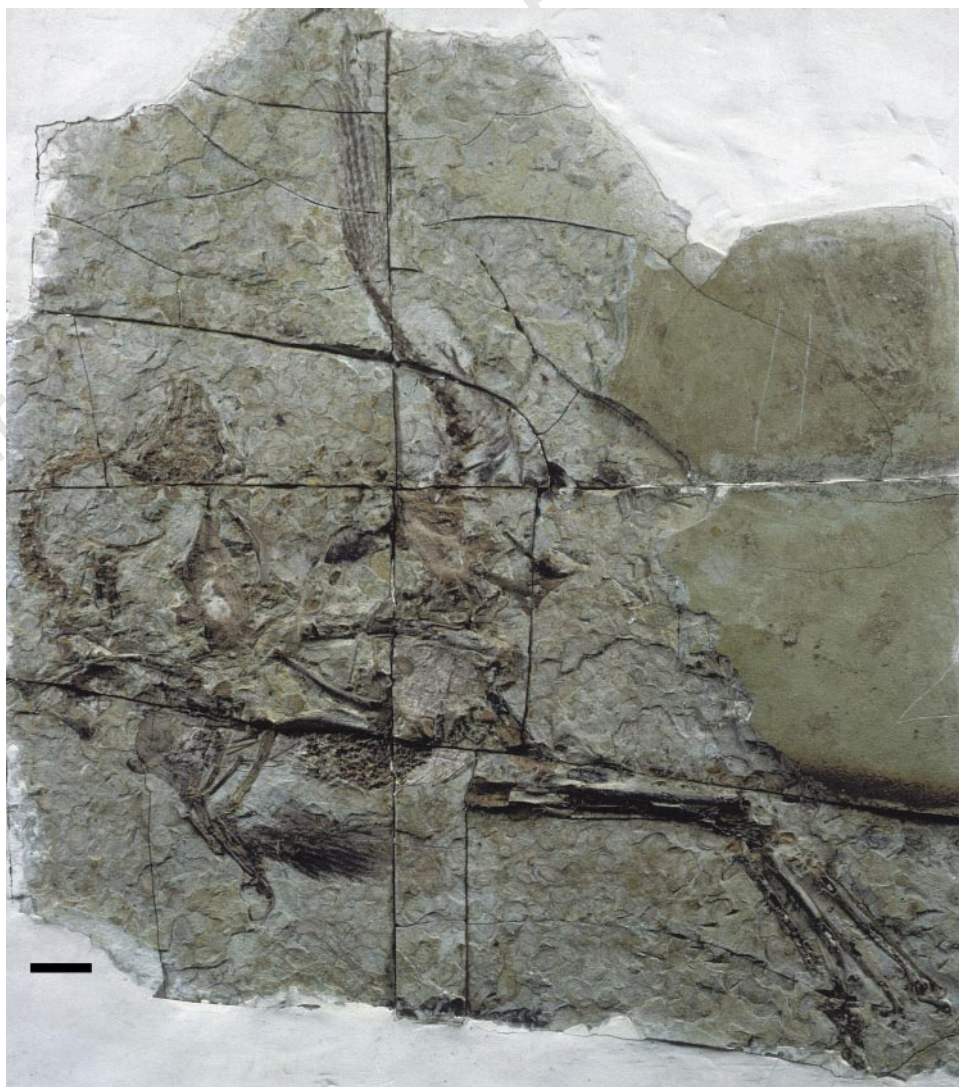


Figure 4 *Caudipteryx zoui*, holotype, NGMC 97-4-A. Scale bar, 5 cm.

The body of NGMC 97-4-A, especially the hips and the base of the tail, is covered by small, plumulaceous feathers of up to 14 mm long.

Both specimens have concentrations of small polished and rounded pebbles in the stomach region. These gastroliths are up to 4.5 mm in diameter, although most are considerably less than 4 mm wide.

Phylogenetic analysis

We examined the systematic positions of *Protarchaeopteryx* and *Caudipteryx* by coding these specimens for the 90 characters used in an analysis of avialan phylogeny²⁴ (for a matrix of these characters, see Supplementary Information). Characters were unordered, and a tree was produced using the branch-and-bound option of PAUP²⁵. We rooted the tree with Velociraptorinae^{26,27}. A single tree resulted with a length of 110 steps, a retention index of 0.849 and a consistency index of 0.855. Analysis shows *Caudipteryx* to be the sister group to the Avialae, and *Protarchaeopteryx* to be unresolved from the Velociraptorinae root (Fig. 9). The placement of *Protarchaeopteryx* as the sister group to *Caudipteryx* + Avialae, as the sister group to Velociraptorinae, or as the sister group to Velociraptorinae + (*Caudipteryx* + Avialae) are equally well supported by the data. Characters that define the *Caudipteryx* + Avialae clade in the shortest tree include unambiguous (uninfluenced by missing data or optimization) characters 2 and 12 and several more

ambiguous ones (characters 4, 5, 10, 11, 15, 19, 24, 37, 85 and 86). *Caudipteryx* is separated from the Avialae by three unambiguous characters (7, 8 and 71) and additional ambiguous ones (characters 5, 6, 9, 10, 11, 18, 24, 39, 40, 56 and 69). The important characteristic of this phylogeny is that the Avialae (not including *Protarchaeopteryx* and *Caudipteryx*) is monophyletic; this placement is supported by the unequivocal presence of a quadratojugal that is joined to the quadrate by a ligament¹⁷ (character 7), the absence of a quadratojugal squamosal contact (character 8) and a reduced or absent process of the ischium (character 71).

As characters dealing with feathers cannot be scored relative to outgroup conditions, they were not used in the phylogenetic analysis. However, our analysis indicates that feathers can no longer be used in the diagnosis of the Avialae.

Discussion

The three *Protarchaeopteryx* and *Caudipteryx* individuals were close to maturity at the time of death. The neural spines seem to be fused to cervical and dorsal centra in *Protarchaeopteryx*. Sternal plates ossify late in the ontogeny of non-avian theropods, and are present in both *Caudipteryx* and *Protarchaeopteryx*. Well-ossified sternal ribs, wrist bones and ankle bones in *Caudipteryx* also indicate the maturity of the specimens.

The remiges of *Caudipteryx* and the rectrices of both

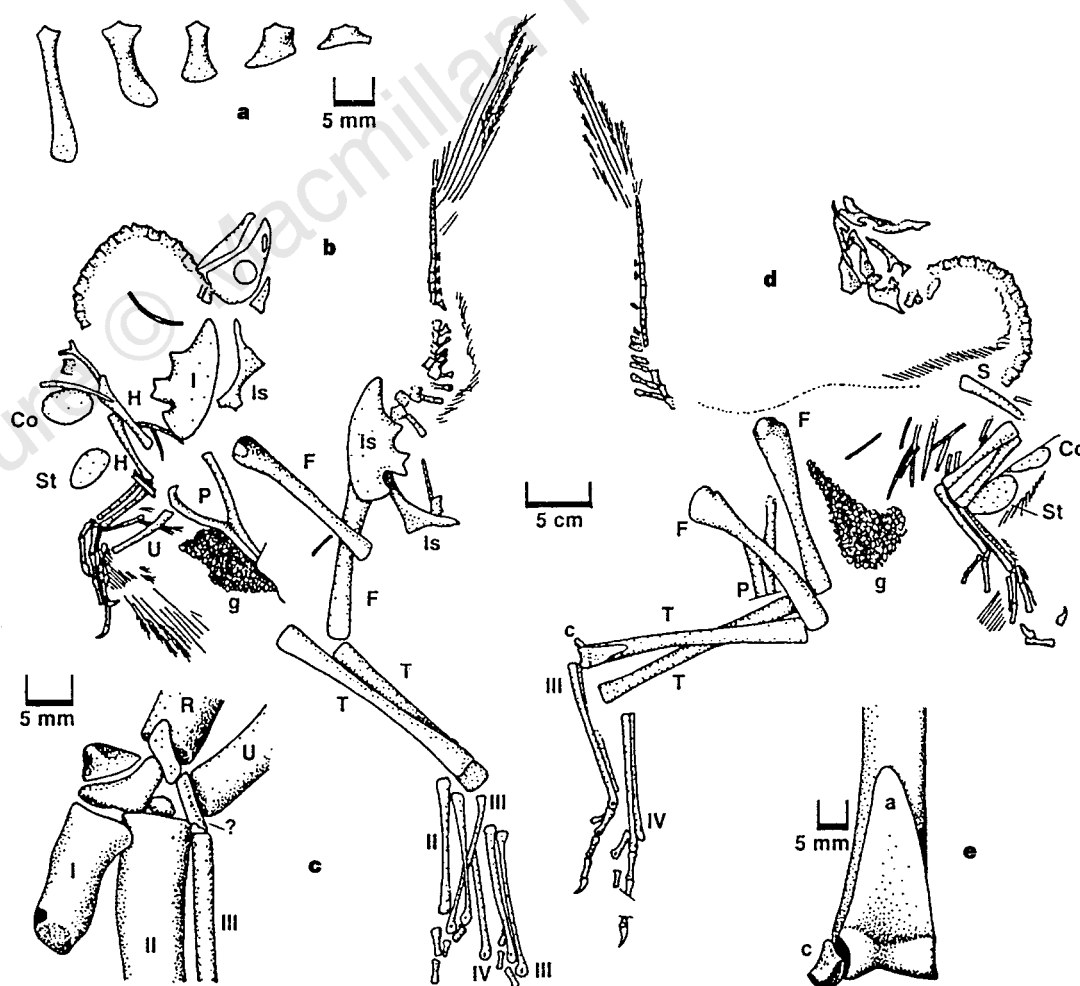


Figure 5 *Caudipteryx zoui*. **a**, Haemal spines from the fourth, sixth, eighth, eleventh and thirteenth caudal vertebrae (from left to right) of NGMC 97-4-A in left lateral view. **b**, Drawing of the specimen shown in Fig. 4a. **c**, Wrist of NGMC 97-4-A. **d**, Drawing of NGMC 97-9-A. **e**, Proximal tarsals of NGMC 97-9-A. Abbrevia-

tions: a, astragalus; c, calcaneum; Co, coracoid; F, femur; g, gastroliths; H, humerus; I, ilium; Is, ischium; P, pubis; R, radius; S, scapula; St, sternal plate; T, tibia; U, ulna; ?, possibly fragment of gastralia. Roman numerals represent digit numbers.



Figure 8 Feathers of *Caudipteryx zoui*, NGMC 97-4-A. **a**, Remiges of left arm. Scale bar, 1.75 cm. **b**, Rectrices, showing colour banding. Scale bar, 1 cm.

Protarchaeopteryx and *Caudipteryx* have symmetrical veins, whereas even those of *Archaeopteryx* are asymmetrical. Birds with asymmetrical feathers are generally considered to be capable of flight²³, but it is possible that an animal with symmetrical feathers could also fly. Relative arm length of *Protarchaeopteryx* is shorter than that of *Archaeopteryx*, but is longer than in non-avian coelurosaurs. The arms of *Caudipteryx*, in contrast, are shorter than those of most non-avian coelurosaurs; the remiges are only slightly longer than the humerus; and the distal remiges are shorter than more proximal

ones. It seems unlikely that this animal was capable of active flight. The relatively long legs of *Protarchaeopteryx* and *Caudipteryx*, both of which have the hallux positioned high and orientated antero-medially, indicate that they were ground-dwelling runners.

Paired rectrices of *Protarchaeopteryx* and *Caudipteryx* are restricted to the end of the tail, whereas in *Archaeopteryx* they extend over more than two-thirds the length of the tail¹². Wherever preservation made it possible, we found semi-plumes and down-like feathers around the periphery of the bodies, suggesting that

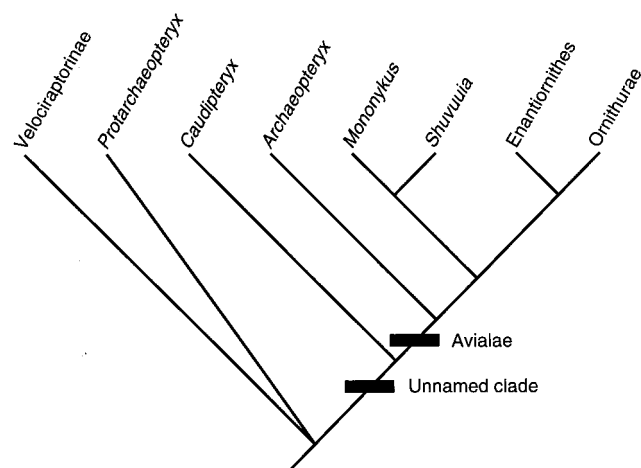


Figure 9 Cladogram of proposed relationships of *Protoarchaeopteryx* and *Caudipteryx*. This tree is based on 90 characters and has a length of 110 steps.

most of the bodies were feather-covered, possibly like *Archaeopteryx*²⁸. Feathers found with *Otogornis*²⁹ were also apparently plumulaceous. Plumulaceous and downy feathers cover the bodies of *Protarchaeopteryx* and *Caudipteryx*, and possibly that of *Sinosauropteryx* as well⁷. This suggests that the original function of feathers was insulation.

Phylogenetic analysis shows that both *Caudipteryx* and *Protarchaeopteryx* lie outside Avialae and are non-avian coelurosaurs. This indicates that feathers are irrelevant in the diagnosis of birds. It can no longer be certain that isolated down and semi-plume feathers^{30–33} discovered in Mesozoic rocks belonged to birds rather than to non-avian dinosaurs. Furthermore, the presence of feathers on flightless theropods suggests that the hypothesis that feathers and flight evolved together is incorrect. Finally, the presence of remiges, rectrices and plumulaceous feathers on non-avian theropods provides unambiguous evidence supporting the theory that birds are the direct descendants of theropod dinosaurs. □

Received 19 January; accepted 27 May 1998.

1. Ji, Q. *et al.* On the sequence and age of the protobird bearing deposits in the Sihetun-Jianshangou area, Beipao, western Liaoning. *Prof. Pap. Strat. Paleol.* (in the press).
2. Hou, L.-H., Zhou, Z.-H., Martin, L. D. & Feduccia, A. A beaked bird from the Jurassic of China. *Nature* **377**, 616–618 (1995).
3. Smith, P. E. *et al.* Dates and rates in ancient lakes: ⁴⁰Ar–³⁹Ar evidence for an Early Cretaceous age for the Jehol Group, northeast China. *Can. J. Earth Sci.* **32**, 1426–1431 (1995).
4. Smith, J. B., Hailu, Y. & Dodson, P. in *The Dinosaur Symposium, Abstracts* (eds Wolberg, D. L. *et al.*) 55 (Academy of Natural Sciences, Philadelphia, 1998).

5. Colbert, E. H. The Triassic dinosaur *Coelophysis*. *Bull. Mus. N. Arizona* **57**, 1–160 (1989).
6. Ostrom, J. H. The osteology of *Compsognathus longipes* Wagner. *Zitteliana* **4**, 73–118 (1978).
7. Chen, P.-j., Dong, Z.-m. & Zhen, S.-n. An exceptionally well-preserved theropod dinosaur from the Yixian Formation of China. *Nature* **391**, 147–152 (1998).
8. Wellnhofer, P. A new specimen of *Archaeopteryx* from the Solnhofen Limestone. *Nat. Hist. Mus. Los Angeles County Sci. Ser.* **36**, 3–23 (1992).
9. Sereno, P. C. & Rao, C. G. Early evolution of avian flight and perching: new evidence from the Lower Cretaceous of China. *Science* **255**, 845–848 (1992).
10. Zhou, Z. H. in *Sixth Symposium on Mesozoic Terrestrial Ecosystems and Biota, Short Papers* (eds Sun, A. & Wang, Y.) 209–214 (China Ocean, Beijing, 1995).
11. Ji, Q. & Ji, S. A. *Protarchaeopteryx* gen. nov.—fossil remains of archaeopterygids from China. *Chinese Geol.* **238**, 38–41 (1997).
12. Wellnhofer, P. Das fünfte skelettexemplar von *Archaeopteryx*. *Palaeontogr. A* **147**, 169–216 (1974).
13. Wellnhofer, P. Das siebte Exemplar von *Archaeopteryx* aus den Solnhofener Schichten. *Archaeopteryx* **11**, 1–48 (1993).
14. Hou, L. H. A carinate bird from the Upper Jurassic of western Liaoning, China. *Chinese Sci. Bull.* **42**, 413–416 (1997).
15. Dong, Z. M. A lower Cretaceous enantiornithine bird from the Ordos Basin of Inner Mongolia, People's Republic of China. *Can. J. Earth Sci.* **30**, 2177–2179 (1993).
16. Forster, C. A., Sampson, S. D., Chiappe, L. M. & Krause, D. W. The theropod ancestry of birds: new evidence from the Late Cretaceous of Madagascar. *Science* **279**, 1915–1919 (1998).
17. Chiappe, L. M., Norell, M. A. & Clark, J. M. The skull of a relative of the stem-group bird *Mononykus*. *Nature* **392**, 275–278 (1998).
18. Norell, M. A. & Makovicky, P. Important features of the dromaeosaur skeleton: information from a new specimen. *Am. Mus. Novit.* **3215**, 1–28 (1997).
19. Perle, A., Chiappe, L. M., Barsbold, R., Clark, J. M. & Norell, M. Skeletal morphology of *Mononykus olecranus* (Theropoda: Avialae) from the Late Cretaceous of Mongolia. *Am. Mus. Novit.* **3105**, 1–29 (1994).
20. Zhou, Z. H. The discovery of Early Cretaceous birds in China. *Courier Forschungsinstitut Senckenberg* **181**, 9–22 (1995).
21. Holtz, T. R. Jr The phylogenetic position of the Tyrannosauridae: implications for theropod systematics. *J. Paleontol.* **68**, 1100–1117 (1994).
22. deBeer, G. *Archaeopteryx lithographica*. *Br. Mus. Nat. Hist.* **244**, 1–68 (1954).
23. Feduccia, A. & Tordoff, H. B. Feathers of *Archaeopteryx*: asymmetric vanes indicate aerodynamic function. *Science* **203**, 1021–1022 (1979).
24. Chiappe, L. M. in *The Encyclopedia of Dinosaurs* (eds Currie, P. J. & Padian, K.) 32–38 (Academic, San Diego, 1997).
25. Swofford, D. & Begle, D. P. *Phylogenetic Analysis Using Parsimony. Version 3.1.1.* (Smithsonian Institution, Washington DC, 1993).
26. Gauthier, J. in *The Origin of Birds and the Evolution of Flight* (ed. Padian, K.) 1–55 (Calif. Acad. Sci., San Francisco, 1986).
27. Holtz, T. R. Jr Phylogenetic taxonomy of the Coelurosauria (Dinosauria: Theropoda). *J. Paleontol.* **70**, 536–538 (1996).
28. Owen, R. On the *Archaeopteryx* of von Meyer, with a description of the fossil remains of a long-tailed species, from the Lithographic Stone of Solenhofen. *Phil. Trans., Lond.* **153**, 33–47 (1863).
29. Hou, L. H. A late Mesozoic bird from Inner Mongolia. *Vert. Palasiatica* **32**, 258–266 (1994).
30. Kurochkin, E. N. A true carinate bird from Lower Cretaceous deposits in Mongolia and other evidence of Early Cretaceous birds in Asia. *Cretaceous Res.* **6**, 271–278 (1985).
31. Sanz, J. L., Bonaparte, J. F. & Lacasa, A. Unusual Early Cretaceous birds from Spain. *Nature* **331**, 433–435 (1988).
32. Kellner, A. W. A., Maisey, J. G. & Campos, D. A. Fossil down feather from the Lower Cretaceous of Brazil. *Palaeontol.* **37**, 489–492 (1994).
33. Grimaldi, D. & Case, G. R. A feather in amber from the Upper Cretaceous of New Jersey. *Am. Mus. Novit.* **3126**, 1–6 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank A. Brush, B. Creisler, M. Ellison, W.-D. Heinrich, N. Jacobsen, E. and R. Koppelhus, P. Makovicky, A. Milner, G. Olshevsky, J. Ostrom and H.-P. Schultze for advice, access to collections and logistic support; and the National Geographic Society, National Science Foundation (USA), the American Museum of Natural History, National Natural Science Foundation of China and the Ministry of Geology for support. Photographs were taken by O. L. Mazzatenta and K. Aulenback; the latter was also responsible for preliminary preparation of the *Caudipteryx* specimens. Line drawings are by P.J.C.

Correspondence and requests for materials should be addressed to P.J.C. (e-mail: pcurrie@mcid.gov.ab.ca).

Tests of quantum gravity from observations of γ -ray bursts

G. Amelino-Camelia^{*†}, John Ellis[‡], N. E. Mavromatos^{*},
D. V. Nanopoulos[§] & Subir Sarkar^{*}

^{*}Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK

[†]Institut de Physique, Université de Neuchâtel, CH-2000 Neuchâtel, Switzerland

[‡]Theory Division, CERN, CH-1211 Geneva, Switzerland

[§]Academy of Athens, Chair of Theoretical Physics, Division of Natural Sciences, 28 Panepistimiou Avenue, Athens GR-10679, Greece; Center for Theoretical Physics, Department of Physics, Texas A & M University, College Station, Texas 77843-4242, USA; and Astroparticle Physics Group, Houston Advanced Research Center (HARC), The Mitchell Campus, Woodlands, Texas 77381, USA

The recent confirmation that at least some γ -ray bursts originate at cosmological distances^{1–4} suggests that the radiation from them could be used to probe some of the fundamental laws of physics. Here we show that γ -ray bursts will be sensitive to an energy dispersion predicted by some approaches to quantum gravity. Many of the bursts have structure on relatively rapid timescales⁵, which means that in principle it is possible to look for energy-dependent dispersion of the radiation, manifested in the arrival times of the photons, if several different energy bands are observed simultaneously. A simple estimate indicates that, because of their high energies and distant origin, observations of these bursts should be sensitive to a dispersion scale that is comparable to the Planck energy scale ($\sim 10^{19}$ GeV), which is sufficient to test theories of quantum gravity. Such observations are already possible using existing γ -ray burst detectors.

Our interest is in the search for possible *in vacuo* dispersion, $\delta v \approx E/E_{\text{QG}}$, of electromagnetic radiation from γ -ray bursts (GRBs), which could be sensitive to a type of candidate quantum-gravity effect that has been recently considered in the particle-physics literature. (Here E is the photon energy and E_{QG} is an effective quantum-gravity energy scale). This candidate quantum-gravity effect would be induced by a deformed dispersion relation for photons of the form $c^2 \mathbf{p}^2 = E^2 [1 + f(E/E_{\text{QG}})]$, where f is a model-dependent function of the dimensionless ratio E/E_{QG} , $\mathbf{p}$ is the photon momentum and c is the velocity of light. In quantum-gravity models in which the hamiltonian equation of motion $\dot{x}_i = \partial H / \partial p_i$ is still valid at least approximately, as in the frameworks discussed later, such a deformed dispersion relation would lead to energy-dependent velocities $c + \gamma v$ for massless particles, with implications for all the electromagnetic signals that we receive from astrophysical objects at large distances. At small energies $E \ll E_{\text{QG}}$, we expect that a series expansion of the dispersion relation should be applicable: $c^2 \mathbf{p}^2 = E^2 [1 + \xi E/E_{\text{QG}} + O(E^2/E_{\text{QG}}^2)]$, where $\xi = \pm 1$ is a sign ambiguity that would be fixed in a given dynamical framework. Such a series expansion would correspond to energy-dependent velocities:

$$v = \frac{\partial E}{\partial p} \approx c \left(1 - \xi \frac{E}{E_{\text{QG}}} \right) \quad (1)$$

This type of velocity dispersion results from a picture of the vacuum as a quantum-gravitational ‘medium’, which responds differently to the propagation of particles of different energies and hence velocities. This is analogous to propagation through a conventional medium such as an electromagnetic plasma⁶. The gravitational ‘medium’ is generally believed to contain microscopic quantum fluctuations, which may occur on scale sizes of order the Planck length $L_p \approx 10^{-33}$ cm on timescales of the order of $t_p \approx 1/E_p$, where $E_p \approx 10^{19}$ GeV. These may^{7,8} be analogous to the thermal fluctuations in a plasma, that occur on timescales of the order of $t \approx 1/T$, where T is the temperature. As it is a much ‘harder’ phenomenon associated with new physics at an energy scale far beyond typical

photon energies, any analogous quantum-gravity effect could be distinguished by its different energy dependence: the quantum-gravity effect would increase with energy, whereas conventional medium effects decrease with energy in the range of interest⁶.

Equation (1) encodes a minute modification for most practical purposes, as E_{QG} is believed to be a very high scale, presumably of the order of the Planck scale $E_p \approx 10^{19}$ GeV. Even so, such a deformation could be rather significant for even moderate-energy signals, if they travel over very long distances. According to equation (1), a signal of energy E that travels a distance L acquires a ‘time delay’, measured with respect to the ordinary case of an energy-independent speed c for massless particles:

$$\Delta t \approx \xi \frac{E}{E_{\text{QG}}} \frac{L}{c} \quad (2)$$

This is most likely to be observable when E and L are large while the interval δt , over which the signal exhibits time structure, is small. This is the case for GRBs, which is why they offer particularly good prospects for such measurements, as we discuss later.

We first review briefly how modified laws for the propagation of particles have emerged independently in different quantum-gravity approaches. The suggestion that quantum-gravitational fluctuations might modify particle propagation in an observable way can already be found in refs 7 and 9. A phenomenological parametrization of the way this could affect the neutral kaon system^{9–11} has been already tested in laboratory experiments, which have set lower limits on parameters analogous to the E_{QG} introduced above at levels comparable to E_p (ref. 12). In the case of massless particles such as the photon, which interests us here, the first example of a quantum-gravitational medium effect with which we are familiar occurred in a string formulation of an expanding Robertson–Walker–Friedman cosmology¹³, in which photon propagation appears tachyonic. Deformed dispersion relations that are consistent with the specific formula in equation (1) arose in approaches based on quantum deformations of Poincaré symmetries¹⁴ with a dimensional parameter. Within this general class of deformations, one finds^{14,15} an effect consistent with equation (1) if the deformation is rotationally invariant: the dispersion relation for massless particles $c^2 \mathbf{p}^2 = E_{\text{QG}}^2 [1 - \exp(E/E_{\text{QG}})]^2$, and therefore $\xi = 1$. We noted that a deformed dispersion relation has also been found in studies of the quantization of point particles in a discrete space time¹⁶.

A specific and general dynamical framework for the emergence of the velocity law (equation (1)) has emerged¹⁷ within the Liouville string approach⁷ to quantum gravity, according to which the vacuum is viewed as a non-trivial medium containing ‘foamy’ quantum-gravity fluctuations. The nature of this foamy vacuum may be visualized by imagining processes that include the pair creation of virtual black holes. Within this approach, it is possible to verify that massless particles of different energies excite vacuum fluctuations differently as they propagate through the quantum-gravity medium, giving rise to a non-trivial dispersion relation of Lorentz ‘non-covariant’ form, just as in a thermal medium. The form of the dispersion relation is not known exactly, but its structure has been studied¹⁷ via a perturbative expansion, and it was shown in ref. 17 that the leading $1/E_{\text{QG}}$ correction is in agreement with equation (1).

It has been recently suggested⁸ the vacuum might have analogous ‘thermal’ properties in a large class of quantum-gravity approaches, namely all approaches in which a minimum length L_{min} —such as the Planck length $L_p \approx 10^{-33}$ cm—characterizes short-distance physics. These should in general lead to deformed photon dispersion relations with $E_{\text{QG}} \approx 1/L_{\text{min}}$, though the specific form of equation (1) may not hold in all models, and hence may be used to discriminate between them. In support of equation (1), though, we recall^{15,17} that this type of non-trivial dispersion in the quantum-gravity vacuum has implications for the measurability of distances in quantum gravity that fit well with the intuition emerging from

recent heuristic analyses¹⁸ based on a combination of arguments from ordinary quantum mechanics and general relativity.

We now explain how GRBs provide an excellent way of testing such ideas, now that the cosmological origin of at least some of them has been established. We recall that typical photon energies in GRB emissions are⁵ in the range 0.1–100 MeV, and it is possible that the spectrum might in fact extend up to TeV energies¹⁹. Moreover, time structure down to the millisecond scale has been observed in the light curves⁵, as is predicted in the most popular theoretical models²⁰ involving merging neutron stars or black holes, where the last stages occur on the timescales associated with grazing orbits. Similar timescales could also occur in models that identify GRBs with other cataclysmic stellar events, such as failed supernovae of type Ib, young ultra-magnetized pulsars or the sudden deaths of massive stars²¹. We see from equations (1) and (2) that a signal with millisecond time structure in photons of energy ~ 20 MeV coming from a distance of the order of 10^{10} light yr, which is well within the range of GRB observations and models, would be sensitive to E_{QG} of the order of 10^{19} GeV $\approx 1/L_{\text{p}}$.

Significant sensitivities may already be attainable with the present GRB data. Submillisecond time-structure has been seen²² in GRB 910711, and a recent time-series analysis²³ of the light curve of GRB 920229 using the bayesian block technique has identified a narrow microburst with a rise and decay timescale of the order of 100 μ s. This is seen simultaneously in three (of the available four) energy channels of the BATSE detector on board the Compton Gamma Ray Observatory, covering the energy regions 20–50 keV, 50–100 keV and 100–300 keV, respectively. From the time structure of this microburst we think it should be possible to extract an upper limit of $\Delta t \leq 10^{-2}$ s on the difference in the arrival times of the burst at energies separated by $\Delta E \approx 200$ keV. If a burst such as this were to be demonstrated in the future to lie at a redshift $z \approx 1$, as seems quite plausible, the implied sensitivity would be to $E_{\text{QG}} \approx 10^{16}$ GeV, and it would be possible to improve this to $\sim 10^{17}$ GeV if the time difference could be brought down to the rise time reported in ref. 23. We note in passing that the simultaneous arrival of photons of different energies from such a large distance also imposes an upper limit of the order of 10^{-6} eV on a possible photon mass, but this is much less stringent than other astrophysical and laboratory limits²⁴.

These levels of sensitivity are even more interesting in light of the fact that recent theoretical work on quantum gravity, particularly within string theory, appears to favour values of the effective scale characterizing the onset of significant quantum-gravity effects that are somewhat below the Planck scale, typically in the range 10^{16} – 10^{18} GeV (ref. 25). If our scale E_{QG} were indeed to be given by such an effective quantum-gravity scale, parts of the GRB spectrum with energies around 0.1 MeV and millisecond time structure (or energies of the order of 100 MeV and 1-s time structure, or energies around 1 TeV and 1-h time structure) might be sensitive to the type of candidate quantum-gravity phenomenon discussed here.

To provide some quantitative comparison of the GRB sensitivity to this phenomenon with that of other astrophysical phenomena, we can compare values of the 'sensitivity factor' $\eta \equiv |\Delta t^*|/\delta t$, where δt represents the time structure of the signal, and Δt^* is the time delay acquired by the signal if $E_{\text{QG}} \approx E_{\text{p}}$, namely $\Delta t^* \approx \pm EL/(cE_{\text{p}})$. As already discussed, GRB emission with millisecond time structure and energy around 20 MeV that travels a distance of the order of 10^{10} light yr has $\eta \approx 1$. Another interesting possibility is that we may observe lensing of a GRB by a foreground galaxy^{26,27}. The burst would then reach us by two or more different paths whose light travel times would differ typically by weeks to years. As conventional gravitational lensing is achromatic, any energy-dependence in the time delay would be a direct probe of the new physics of interest, and would be independent of the actual emission mechanism of γ -ray bursts. We note that the HEGRA²⁸ and Whipple²⁹ air Čerenkov telescopes have already searched for TeV emission from the direction

of GRBs, motivated by the EGRET telescope's detection³⁰ of emission up to 18 GeV from GRB940217. If such searches were to prove successful and moreover, identified a lensed GRB, one would be able to infer via equation (2) a sensitivity factor down to $\eta \approx 10^{-6}$.

As observed in ref. 17, which did not consider GRBs as the cosmological distances of these phenomena were not then established, pulsars and supernovae are among the other astrophysical phenomena that might at first sight appear well suited for probing the physics we are interested in here, because of the short time structures they display. However, although pulsar signals have very well-defined time structure, they are at relatively low energies and are observable over distances of at most 10^4 light yr. If one takes an energy of the order of 1 eV and postulates a sensitivity to time delays as small as 1 μ s, one estimates a sensitivity down to $\eta \approx 10^{-10}$. With new experiments such as AXAF it may be possible to detect X-ray pulsars out to 10^6 light years, allowing us to probe down to $\eta \approx 10^{-8}$.

We note that neutrinos from type-II supernovae like SN1987a—which should have energies up to ~ 100 MeV with a time structure that could extend down to milliseconds—are likely to be detectable at distances of up to $\sim 10^5$ light yr, providing sensitivity down to $\eta \approx 10^{-4}$. We have also considered the cosmic microwave background. Although the distance travelled by these photons is the largest available, the only possible signature is a small distortion of the Planck spectrum due to the frequency-dependence of c , which we estimate to be of the order of $\Delta I(\nu)/I(\nu) \approx \nu/E_{\text{p}} \approx 10^{-32}$, where $I(\nu)$ is the frequency spectrum, which is negligible.

In principle, GRBs allow us to gain many orders of magnitude in the sensitivity factor η . Moreover, and most importantly, this high sensitivity should be sufficient to probe values of the effective scale characterizing the onset of quantum-gravity effects extending all the way up to the Planck scale, as illustrated by the estimates we have provided. Ideally, one would like to understand well the short-time structure of GRB signals in terms of conventional physics, so that the phenomena discussed here may be disentangled unambiguously. However, even in the absence of a complete theoretical understanding, sensitive tests can be performed as indicated above, through the serendipitous detection of short-scale time structure²³ at different energies in GRBs which are established to be at cosmological distances. Detailed features of burst time series should enable the emission times in different energy ranges to be put into correspondence. As any time shift due to quantum-gravity effects of the type discussed here would increase with the photon energy, this characteristic dependence should be separable from more conventional in-medium physics effects, which decrease with energy. To distinguish any quantum-gravity shift from effects due to the source, we recall that the medium effect would be linear in the source distance and in the photon energy, which would not in general be the case for time shifts at the source. To disentangle any such effects, it is clear that the most desirable features in an observational programme would be fine time resolution at high photon energies. $\square$

Received 12 December 1997; accepted 10 April 1998.

1. van Paradijs, J. *et al.* Transient optical emission from the error box of the γ -ray burst of 28 February 1997. *Nature* **386**, 686–689 (1997).
2. Groot, P. J. *et al.* IAU Circ. No. 6676 (1997).
3. Metzger, M. L. *et al.* Spectral constraints on the redshift of the optical counterpart to the γ -ray burst of 8 May 1997. *Nature* **387**, 878–880 (1997).
4. Metzger, M. L. *et al.* IAU Circ. No. 6676 (1997).
5. Fishman, G. J. & Meegan, C. A. Gamma-ray bursts. *Annu. Rev. Astron. Astrophys.* **33**, 415–458 (1995).
6. Latorre, J. L., Pascual, P. & Tarrach, R. Speed of light in nontrivial vacua. *Nucl. Phys. B* **437**, 60–82 (1995).
7. Ellis, J., Mavromatos, N. & Nanopoulos, D. V. String theory modifies quantum mechanics. *Phys. Lett. B* **293**, 37–48 (1992).
8. Garay, L. J. Space-time foam as a quantum thermal bath. *Phys. Rev. Lett.* (submitted); also as preprint gr-qc/9801024.
9. Ellis, J., Hagelin, J. S., Nanopoulos, D. V. & Srednicki, M. Search for violations of quantum mechanics. *Nucl. Phys. B* **241**, 381–405 (1984).
10. Ellis, J., Lopez, J., Mavromatos, N. & Nanopoulos, D. V. Precision tests of CPT symmetry and quantum mechanics in the neutral kaon system. *Phys. Rev. D* **53**, 3846–3870 (1996).
11. Huet, P. & Peskin, M. E. Violation of CPT and quantum mechanics in the $K_0 - \bar{K}_0$ system. *Nucl. Phys. B* **434**, 3–38 (1995).
12. Adler, R. *et al.* Tests of CPT symmetry and quantum mechanics with experimental data from CPLEAR. *Phys. Lett. B* **364**, 239–245 (1995).

13. Antoniadis, I., Bachas, C., Ellis, J. & Nanopoulos, D. V. Comments on cosmological string solutions. *Phys. Lett. B* **257**, 278–284 (1991).
14. Lukierski, J., Nowicki, A. & Ruegg, H. Classical and quantum-mechanics of free κ -relativistic systems. *Ann. Phys.* **243**, 90–116 (1995).
15. Amelino-Camelia, G. Enlarged bound on the measurability of distances and quantum κ -Poincaré group. *Phys. Lett. B* **392**, 283–286 (1997).
16. 't Hooft, G. Quantization of point particles in $(2 + 1)$ -dimensional gravity and space-time discreteness. *Class. Quant. Grav.* **13**, 1023–1039 (1996).
17. Amelino-Camelia, G., Ellis, J., Mavromatos, N. E. & Nanopoulos, D. V. Distance measurement and wave dispersion in a Liouville string approach to quantum gravity. *Int. J. Mod. Phys. A* **12**, 607–623 (1997).
18. Amelino-Camelia, G. Limits on the measurability of space-time distances in the semi-classical approximation of quantum gravity. *Mod. Phys. Lett. A* **9**, 3415–3422 (1994).
19. Baring, M. G. Gamma-ray bursts above 1 GeV. in *Towards a Major Atmospheric Cerenkov Detector* (ed. de Jager, O. C.) (Proc. Kruger National Park TeV Workshop, Westprint, Potchefstroom, in the press); also as preprint astro-ph/9711256.
20. Rees, M. J. Gamma-ray bursts: challenges to relativistic astrophysics. in *Proc. 18th Texas Symp. on Relativistic Astrophysics 1996* (eds Olinto, A., Friemann, J. & Schramm, D. N.) (World Scientific, in the press); also as preprint astro-ph/9701162.
21. Mészáros, P. Theoretical models of gamma-ray bursts. in *Gamma-Ray Bursts* (eds Meegan, C., Preece, R. & Koshut, T.) (Proc. 4th Huntsville Symp., Am. Inst. Phys., in the press); also as preprint astro-ph/9711354.
22. Bhat, C. L. *et al.* Evidence for sub-millisecond structure in a γ -ray burster. *Nature* **359**, 217–216 (1992).
23. Scargle, J. D., Norris, J. & Bonnell, J. Attributes of GRB pulses: Bayesian blocks analysis of TTE data; a microburst in GR920229. in *Gamma-Ray Bursts* (eds Meegan, C., Preece, R. & Koshut, T.) (Proc. 4th Huntsville Symp., Am. Inst. Phys., in the press); also as preprint astro-ph/9712016.
24. Barnett, R. M. *et al.* Review of particle properties. *Phys. Rev. D* **54**, 207–720 (1996).
25. Witten, E. Strong coupling expansion of Calabi-Yau compactification. *Nucl. Phys. B* **471**, 135–158 (1996).
26. Paczynski, B. Gamma-ray bursts at cosmological distances. *Astrophys. J.* **308**, L43–L46 (1986).
27. Nemiroff, R. J. *et al.* Searching gamma-ray bursts for gravitational lensing echoes—Implications for compact dark matter. *Astrophys. J.* **414**, 36–40 (1994).
28. Krawczynski, H. *et al.* Search for TeV counterparts of gamma-ray bursts with the HEGRA experiment. in *Proc. Int. School of Cosmic-Ray Astrophysics* (World Scientific, in the press); also as preprint astro-ph/9611044.
29. Boyle, P. J. *et al.* in *Proc. 25th Int. Cosmic ray Conf. Vol. 3* (eds Potgieter M. S. *et al.*) 61 (Westprint, Potchefstroom, 1998); also as preprint astro-ph/9706132.
30. Hurley, K. *et al.* Detection of a γ -ray burst of very long duration and very high energy. *Nature* **372**, 652–654 (1994).

Acknowledgements. G.A.-C. thanks J. Binney and other members of the Oxford Theoretical Astrophysics group for discussions. This work was supported in part by the Foundation Blancefort Boncompagni-Ludovisi (G.A.-C.), a PPARC advanced fellowship (N.E.M.) and the US Department of Energy (D.V.N.)

Correspondence and requests for materials should be addressed to G.A.-C. at the University of Neuchâtel (e-mail: Giovanni.Amelino-Camelia@cern.ch).

Global warming on Triton

J. L. Elliot^{*†‡}, H. B. Hammel^{*}, L. H. Wasserman[‡],
O. G. Franz[‡], S. W. McDonald^{*}, M. J. Person^{*}, C. B. Olkin[‡],
E. W. Dunham[‡], J. R. Spencer[‡], J. A. Stansberry[‡],
M. W. Buie[‡], J. M. Pasachoff[§], B. A. Babcock^{||}
& T. H. McConnochie[§]

^{*} Department of Earth, Atmospheric, and Planetary Sciences, [†] Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA

[‡] Lowell Observatory, Flagstaff, Arizona 86001, USA

[§] Astronomy Department, ^{||} Physics Department, Williams College, Williamstown, Massachusetts 01267-2565, USA

Triton, Neptune's largest moon, has been predicted to undergo significant seasonal changes that would reveal themselves as changes in its mean frost temperature^{1–3}. But whether this temperature should at the present time be increasing, decreasing or constant depends on a number of parameters (such as the thermal properties of the surface, and frost migration patterns) that are unknown. Here we report observations of a recent stellar occultation by Triton which, when combined with earlier results, show that Triton has undergone a period of global warming since 1989. Our most conservative estimates of the rate of temperature and surface-pressure increase during this period imply that the atmosphere is doubling in bulk every 10 years—significantly faster than predicted by any published frost model for Triton^{2,3}. Our result suggests that permanent polar caps on Triton play a dominant role in regulating seasonal atmospheric changes. Similar processes should also be active on Pluto.

The 4 November 1997 occultation of the star Tr180 (also known

as Tycho 651672 and GSC6321–01030) was successfully observed with the Hubble Space Telescope (HST) in daylight over the northwest Pacific Ocean; Astrometer 3 of the Fine Guidance Sensors (FGS) was used to record the event⁵. Details of the HST data are given in Table 1 along with information about other observing stations; the immersion and emersion data (disappearance and reappearance of the star) are shown in Fig. 1. A central flash (the focusing of light rays by Triton's atmosphere^{6,7}) was recorded, but will be presented and analysed elsewhere.

We modelled the HST light curve with a standard small-planet model that allows for a power-law temperature gradient⁸ (Table 2). The background from dark counts and Triton (determined by the FGS in September and adjusted for Triton's different distance) was subtracted, and the remainder was divided by the flux from the star (also determined by the FGS in September) so that the full range of stellar flux corresponded to values between 0.0 and 1.0. In the light-curve model fits, the zero level was fixed, but the full-scale signal from the star was a free parameter. The difference of the fitted values from 1.0 shows that our calibration error is only a few tenths of one per cent.

From fitting the entire light curve, the closest-approach distance between the centre of Triton's shadow and the HST was determined to be 224 ± 4 km (first column of results in Table 2). This value places the shadow somewhat further north than predicted, but it is consistent with no detectable occultation at our Oahu station (Table 1). Using the closest-approach distance determined from the entire light curve, we fitted the main immersion and emersion sections of the light curve both separately and together (next three columns of results in Table 2). As a test of the self-consistency of our light-curve models, we fixed the atmospheric model parameters ("half-light radius", "lambda at half-light", and the "thermal-gradient exponent"; see ref. 8) at their values determined from

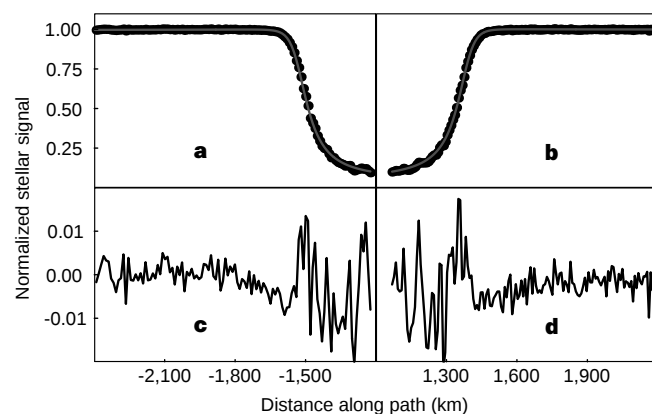


Figure 1 Triton occultation light curves from the HST. Data before and after the occultation were used to establish the modulation of the signal due to the astrometer scan, which was then removed from the entire data set. The zero and full-scale stellar flux levels were established with photometric data from an earlier FGS visit to these objects on 11 September 1997. At that time, Triton's magnitude as observed by the FGS was 13.4 and the magnitude of Tr180 was 10.6. The FGS data have been averaged at 1.0 s; **a**, immersion data (filled circles) and light-curve model fit (line); **b**, emersion data (filled circles) and light-curve model fit (line; see fits in Table 2). The zero point of the abscissa is arbitrary. The light-curve model⁸ used a power-law thermal gradient; residuals from the fit in **a** are shown in **c**, and the residuals from the fit in **b** are shown in **d**. The r.m.s. residual along the full signal is ~ 0.0022 for a 1-s average, and is the result of photon noise. The remaining residuals that occur when the star is partially occulted are the result of unmodelled structure in Triton's atmosphere and rarely exceed 0.01. The effect of these on our determination of the 1,400-km pressure can be estimated by differences among the fits in Table 1. The omitted central portion of the light curve (see text) corresponds to radii $< 1,400$ km (altitudes < 48 km); the atmosphere below that does not affect the 1,400-km pressure.

the combined immersion and emersion fit (second to last column in Table 2), and then fit the entire light curve again (last column in Table 2).

Although the agreement is not perfect between all model fits for the atmospheric parameters at a radius of 1,400 km, it is very good. In particular, the close agreement between the immersion and emersion atmospheric parameters for two widely separated locations on Triton supports the idea that the sublimation and condensation of nitrogen maintains the surface frost at the same temperature. The pressure at 1,400 km, $2.3 \pm 0.1 \mu\text{bar}$, is derived from the fit to the immersion and emersion data (the error bar includes the systematic differences between the fits in Table 1). This is significantly greater than the 1,400-km pressure of $1.4 \pm 0.1 \mu\text{bar}$ measured with stellar occultations in 1995⁹ and the value of $0.8 \pm 0.1 \mu\text{bar}$ extrapolated from a surface pressure of $14 \pm 1 \mu\text{bar}$ measured by Voyager in 1989^{10,11}. The pressure at 1,400 km has certainly increased between 1995 and 1997, based on the two sets of stellar occultation data; the temperature at 1,400 km has also increased from $47 \pm 1 \text{ K}$ (ref. 9) to $50.3 \pm 0.5 \text{ K}$. As these temperatures are consistent with the 1,400-km temperature predicted by atmospheric models based on Voyager data^{12,13}, a surface-pressure increase between the 1989 and more recent measurements is the most likely explanation for the difference (rather than using an inappropriate model for extrapolating the pressure from the surface to 1,400 km).

This surface-pressure increase implies a temperature increase of the surface frost, as the principal constituent of the tenuous atmosphere, nitrogen, is presumed to be in vapour-pressure equilibrium with the surface frost. The surface pressures and corresponding equilibrium temperatures for N_2 are plotted in Fig. 2, where two cases are shown.

In the first case, we have extrapolated the occultation pressures from a radius of 1,400 km to the surface radius of 1,352 km (an average of published values^{14,15}) using a multiplicative factor of 17.5, which is valid under the assumption that the shape of the thermal profile¹² has not significantly changed since the Voyager encounter (filled symbols in Fig. 2). We justify this assumption by the consistency of the atmospheric temperature at 1,400 km (Table 2) with the models^{12,13}. These measurements show a steady increase in pressure and surface-frost temperature. At this rate the atmosphere has tripled in bulk since the time of the Voyager encounter, which would require $\sim 0.9\%$ of the total solar energy incident on Triton during the intervening time to sublime the frost. In the second case, we make a more conservative estimate of the surface pressure derived from the stellar occultation data by assuming an isothermal atmosphere below 1,400 km (open symbols in Fig. 2). Even for this case, the atmosphere has nearly doubled in bulk since the time of Voyager.

Triton's atmosphere is supported by surface frosts, which at any instant are expected to be at the same temperature everywhere, due

Table 1 Observations

Site	Telescope aperture (m)	Detector	Filter	Recording interval (ut on 4 Nov. 1997)	Integration time (s)	Aperture (arcsec)	S/N*
HST†	2.5	4 PMTs‡	F583W	04:59:32–05:24:44	0.025	5 × 5	630
LCC§	0.3	CCD	Nonell	05:10:55–05:27:35	0.125	15 × 15	33

* S/N refers to the signal-to-noise ratio of the occultation light curve (when the star is unocculted) over a time interval that corresponds to 20 km (approximately a scale height) of relative motion between the telescope and Triton's shadow.

† HST was located (in J2000 coordinates, km relative to the centre of the Earth) at $(x, y, z) = (-3,906, -5,752, 761)$ for immersion, and $(x, y, z) = (-1,898, -6,484, 1,806)$ for emersion. Centre-line observations with NASA's Lear Jet Observatory were attempted, but failed due to target acquisition and tracking problems. Simultaneous optical and infrared observations with NASA's Infrared Telescope Facility on Mauna Kea (MKO) and optical observations with a 0.4-m portable telescope on Tern Island were foiled by clouds; other groups had similar bad weather on the University of Hawaii 2.2-m telescope on MKO and on Kauai (more details are provided at <http://occult.mit.edu/tr180plans/>).

‡ The four photomultiplier tubes (PMTs) of Astrometer 3 of the Fine Guidance Sensors were used in "TRANS" mode. We generated the light curve by adding the signals from all four PMTs⁵. § The telescope at Leeward Community College (LCC) is located at longitude $-157^\circ 59' 6'' \text{ E}$, latitude $21^\circ 23' 30'' \text{ N}$, and an altitude of 10 m. The data are a continuous series of CCD frames (each 60 by 57 pixels). The light curve was generated from these frames with synthetic aperture photometry of the blended image of Triton and the star. This station recorded data through the time of the event but saw no obvious occultation, which indicates that Triton's shadow path went somewhat north of the prediction (provided at <http://occult.mit.edu/tr180.html>).

|| The effective wavelength of the F583W filter was 583 nm; the open CCD used at LCC had an effective wavelength of 670 nm.

Table 2 Models for the Triton occultation light curve obtained with the HST

	Data selection				
	All	Immersion*	Emersion*	Immer. & emer.	All
Model parameters					
Background level	0.00	0.00	0.00	0.00	0.00
Background slope (10^{-7} km^{-1})	5.3 ± 0.7	0.00	0.00	2.2 ± 0.8	4.9 ± 0.6
Star signal	0.9968 ± 0.0003	0.9957 ± 0.0004	0.9974 ± 0.0003	0.9966 ± 0.0002	0.9971 ± 0.0003
Closest approach time†	3.828 ± 0.034	3.808	3.808	3.808 ± 0.015	3.817 ± 0.034
Half-light radius (km)	$1,456.3 \pm 0.6$	$1,456.3 \pm 0.3$	$1,456.4 \pm 0.2$	$1,456.3 \pm 0.2$	1,456.3
Lambda (isothermal) at half-light	65.3 ± 0.4	65.5 ± 0.4	66.4 ± 0.3	67.4 ± 0.2	67.4
Thermal-gradient exponent	1.9 ± 0.3	1.7 ± 0.5	4.4 ± 0.4	2.9 ± 0.3	2.9
Minimum centre distance (km)	224 ± 4 ‡	224	224	224	219 ± 2
Derived quantities at $r = 1,400 \text{ km}$ (altitude = 48 km)					
Number density (10^{14} cm^{-3})	3.23 ± 0.09	3.23 ± 0.06	3.42 ± 0.04	3.31 ± 0.04	3.31
Pressure (μbar)	2.25 ± 0.06	2.24 ± 0.04	2.36 ± 0.03	2.30 ± 0.03	2.30
Temperature (K)	50.6 ± 0.3	50.4 ± 0.3	50.1 ± 0.2	50.3 ± 0.2	50.3
Temperature gradient (K km^{-1})	0.07 ± 0.01	0.06 ± 0.02	0.16 ± 0.01	0.10 ± 0.01	0.10
Other derived quantities					
Minimum altitude probed (km)§	<25	46	46	46	<26
Surface pressure (μbar)	39 ± 4	39 ± 4	41 ± 4	40 ± 4	40
Fit information					
Degrees of freedom	1,504	321	321	644	1,507
χ^2 per degree of freedom	17.8	8.94	4.22	6.77	18.1

These models are for a pure N_2 atmosphere. Model parameters with error bars were fitted; those without were held constant.

* The immersion data included the interval from 05:01:03 to 05:06:27 ut on 4 Nov. 1997; the emersion interval was 05:10:06 to 05:15:30 ut. Immersion probed the atmosphere above -5.1° latitude and 10.4° longitude on Triton; emersion probed the atmosphere above -8.4° latitude and 206.4° longitude.

† Seconds after 05:08:00 ut on 4 Nov. 1997.

‡ The minimum centre distance is primarily controlled by the amplitude of the central flash, which will be discussed elsewhere.

§ For the fits to all the data, the 'minimum altitude probed' refers to the light-curve centre. A surface radius of 1,352 km was used^{14,15}.

|| χ^2 per degree of freedom is much greater than 1.0 for all fits, indicating that the deviations of the atmospheric structure from a power-law thermal gradient⁶ (rather than photon noise) are the primary source of error (note the behaviour of the residuals in the lower panels of Fig. 1).

to efficient transfer of latent heat through the atmosphere^{16,17}. This global frost temperature determines atmospheric pressure through vapour-pressure equilibrium; thus a frost temperature increase from 37.5 K (in 1989) to 39.3 K (in 1997) can be inferred from the observed increase in atmospheric pressure (the non-isothermal case in Fig. 2).

The mechanisms that can cause a change in frost temperature (and hence surface pressure) are: (1) non-static (migrating) surface-frost distribution; (2) changes in the optical properties of the surface frost; and (3) changing insolation on a static surface-frost distribution. (Although subsurface heat flow—geothermal or seasonal—can affect the energy balance, it is unlikely to have varied rapidly in the past decade.)

Simple seasonal models of long-term frost migration, with uniform frost albedo and emissivity^{1,3}, show monotonic long-term transfer of most of Triton's nitrogen frost to the poles, which have the minimum seasonally averaged insolation. In most of these models, the remaining seasonal frost migrates between hemispheres in response to the seasonal insolation cycle, and all frosts should be almost exhausted from the southern hemisphere during the current extreme southern summer. This would result in a rapid decrease in atmospheric pressure, the opposite of what our 1995 and 1997 observations indicate has occurred.

A substantial decrease of mean frost albedo or emissivity could account for the pressure change. For instance, model "F" of Spencer and Moore³, which assumes unit emissivity, can account for the atmospheric change if the mean frost Bond albedo dropped by 0.12 between 1989 and 1997. Observational^{18,19} and theoretical^{20,21} work suggests that large changes in frost optical properties may indeed occur on Triton.

The frost's total absorbed insolation may have increased, due to the increase in the southerly subsolar latitude in the past decade. We computed the change in surface pressure that would result solely from the change in subsolar latitude from August 1989 (47.8° south) to November 1997 (51.0° south), under the assumption that no volatile transport had occurred in the intervening 8 years. Five different maps of the 1989 N₂-ice distribution on Triton were generated based on the albedos measured by Voyager. The emissiv-

ity of the N₂ ice was taken as a free parameter in each of these frost models, and was adjusted such that the temperature of the N₂ was equal to 37.5 K (corresponding to 14 μ bar surface pressure) in 1989. The N₂-ice thermal balance was then recomputed using the subsolar latitude appropriate for Triton in late 1997, under the assumption that the distribution of N₂ ice was unchanged (albedos did change slightly, due to the illumination-angle dependence of the bolometric albedo, and the rotation of some N₂-ice-covered portions of the globe into darkness). Several of these models result in pressure increases comparable to the measured increase. Thus, the observed global warming of Triton may be due to increased insolation of a permanent south polar cap. □

Received 12 February; accepted 14 April 1998.

1. Spencer, J. R. Nitrogen frost migration on Triton: a historical model. *Geophys. Res. Lett.* **17**, 1769–1772 (1990).
2. Hansen, C. J. & Paige, D. A. A thermal model for the seasonal nitrogen cycle on Triton. *Icarus* **99**, 273–288 (1992).
3. Spencer, J. R. & Moore, J. M. The influence of thermal inertia on temperatures and frost stability on Triton. *Icarus* **99**, 261–272 (1992).
4. McDonald, S. W. & Elliot, J. L. Triton stellar occultation candidates: 1995–1999. *Astron. J.* **109**, 1352–1362 (1995).
5. Holfelt, S. T. *FGS Instrument Handbook* (Space Telescope Science Institute, Baltimore, 1996).
6. Elliot, J. L. et al. Occultation of ϵ Geminorum by Mars. II. The structure and extinction of the Martian upper atmosphere. *Astrophys. J.* **217**, 661–679 (1977).
7. Elliot, J. L., Stansberry, J. A., Olkin, C. B., Agner, M. A. & Davies, M. E. Triton's distorted atmosphere. *Science* **278**, 436–439 (1997).
8. Elliot, J. L. & Young, L. A. Analysis of stellar occultation data for planetary atmospheres. I. Model fitting, with application to Pluto. *Astron. J.* **103**, 991–1015 (1992).
9. Olkin, C. B. et al. The structure of Triton's atmosphere: Results from the entire ground-based occultation data set. *Icarus* **129**, 178–201 (1997).
10. Tyler, G. L. et al. Voyager radio science observations of Neptune and Triton. *Science* **246**, 1466–1473 (1989).
11. Gurrola, E. M. *Interpretation of Radar Data from the Icy Galilean Satellites and Triton* Thesis, Stanford Univ. (1995).
12. Strobel, D. F., Zhu, X., Summers, M. E. & Stevens, M. H. On the vertical thermal structure of Pluto's atmosphere. *Icarus* **120**, 266–289 (1996).
13. Krasnopolsky, V. A., Sandel, B. R., Herbert, F. & Vervack, R. J. Temperature, N₂, and N density profiles of Triton's atmosphere: observations and model. *J. Geophys. Res.* **98**, 3065–3078 (1993).
14. Davies, M. E., Rogers, P. G. & Colvin, R. R. A control network of Triton. *J. Geophys. Res.* **96**, 15675–15681 (1991).
15. Marouf, E. A., Tyler, G. L., Eshleman, V. R. & Rosen, P. A. Voyager radio occultation of Triton: surface topography and radius. *Bull. Am. Astron. Soc.* **23**, 1207 (1991).
16. Trafton, L. Large seasonal variations on Triton. *Icarus* **58**, 312–324 (1984).
17. Ingersoll, A. P. Dynamics of Triton's atmosphere. *Nature* **344**, 315–317 (1990).
18. Smith, B. A. et al. Voyager 2 at Neptune: Imaging science results. *Science* **246**, 1422–1449 (1989).
19. Buratti, B. J., Goguen, J. D., Gibson, J. & Mosher, J. Historical photometric evidence for volatile migration on Triton. *Icarus* **110**, 303–314 (1994).
20. Eluzkiewicz, J. On the microphysical state of the surface of Triton. *J. Geophys. Res.* **96**, 19217–19230 (1991).
21. Duxbury, N. S. & Brown, R. H. The phase composition of Triton's polar caps. *Science* **261**, 748–751 (1993).

Acknowledgements. We thank E. Nelan and B. McArthur for help in reducing our data; the IRTF staff, the staff of the Lear Jet Observatory at NASA Ames, F. Osell, M. Kakkala, J. Kern, R. Meserole, C. Dahn and R. Stone for assistance; and D. Strobel for discussions. This work was supported, in part, by NASA, NSF, and the National Geographic Society. The NASA/ESA Hubble Space Telescope observations were supported by STScI, which is operated by Association of Universities for Research in Astronomy, Incorporated.

Correspondence and requests for materials should be addressed to J.L.E. at the Department of Earth, Atmospheric, and Planetary Sciences, MIT.

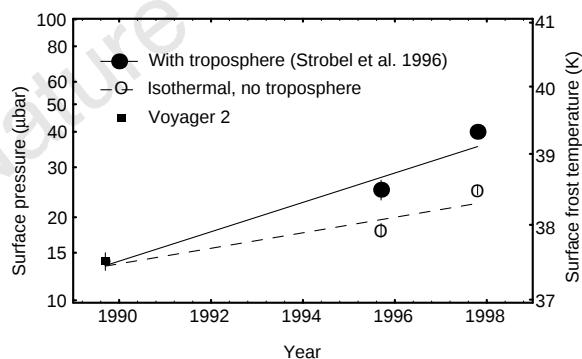


Figure 2 Triton's global warming. The surface pressure and the corresponding vapour-pressure equilibrium temperature is plotted versus year of measurement. The measurement in 1989 refers to the results of an atmospheric model¹² based on several measurements made by Voyager 2. The 1995 data are from stellar occultations of Tr148A and Tr148B³, and the 1997 data refer to the present work. Each occultation datum has been plotted twice along with a corresponding least-squares fit: assuming that the atmosphere follows the model of Strobel et al.¹² (filled circles, solid line), and making the conservative assumption that the atmosphere is isothermal below 1,400 km (open circles, dashed line). For the conservative assumption, the pressure increase is $1.4 \pm 0.4 \mu\text{bar yr}^{-1}$; it is $2.8 \pm 0.8 \mu\text{bar yr}^{-1}$ for the other case. In both cases, the pressure is seen to steadily increase, indicating that Triton has been undergoing a period of global warming. The distance of the points from the lines is larger than their error bars, which indicates that the warming of the surface frost may not have occurred linearly with time.

Superconductivity in oxygen

K. Shimizu^{*†}, K. Suhara^{*}, M. Ikumo^{*}, M. I. Eremets[‡] & K. Amaya^{*†‡}

^{*} Department of Material Physics, Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

[†] CREST of Japan Science and Technology Corporation, Kawaguchi, Saitama 332-0012, Japan

[‡] Advanced Science Research Center, Japan Atomic Energy Research Institute, Naka, Ibaraki 319-1195, Japan

Among the simple diatomic molecules, oxygen is of particular interest because it shows magnetism at low temperatures. Moreover, at pressures exceeding 95 GPa (~ 0.95 Mbar), solid molecular oxygen becomes metallic, accompanied by a structural transition¹. The metallization process is characterized by an increase in optical reflectivity², and a change in the slope of the resistance—

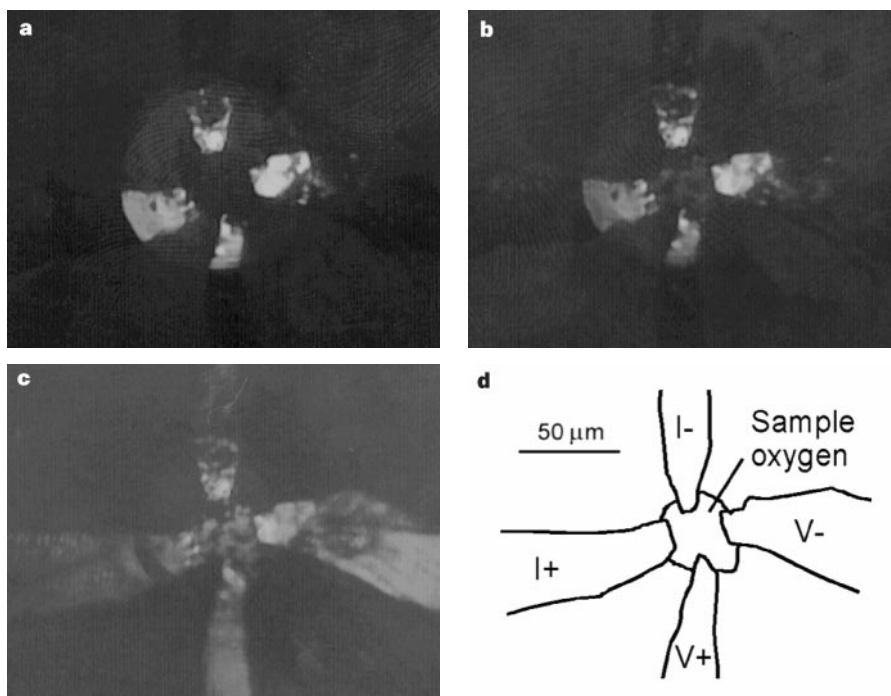


Figure 1 Photographs of the oxygen sample and four electrodes on an insulated gasket in the pressure cell at pressures of **a**, 40; **b**, 60; and **c**, 120 GPa. **c**, Metallic reflection from oxygen at 120 GPa can be compared with the platinum electrodes. **d**, The measuring current is applied through the platinum electrode of I+ to I- and resistance is obtained from the voltage between V+ and V-.

temperature curve³. Here we report that at pressures of around 100 GPa, solid oxygen becomes superconducting, with a transition temperature of 0.6 K. The transition is revealed by both resistivity measurements and a Meissner demagnetization signal.

We used a diamond anvil cell with a pair of 1/3 carat synthetic diamonds with a top surface diameter of 100 μm to produce very high pressures (above 100 GPa). The surface is bevelled by 7 degrees from a culet with a diameter of 300 μm. The pressure cell is made of pure (cobalt-free) Be-Cu alloy, that produced a negligible magnetic effect down to the lowest measured temperature. We filtered oxygen gas (99.998% purity) through a cold trap and liquefied it in a liquid nitrogen cryostat. The oxygen sample was confined by external pressurizing, and the cell was warmed to room temperature and pressurized to the measuring pressures. The pressure value was set at the temperature of liquid nitrogen (77 K) and decreasing the temperature further did not change the pressure. This is because of the small thermal expansion of the cell at the lower temperature region.

Figure 1 shows photographs of the oxygen sample and electrodes

under pressure, which were taken through the diamond window. The appearance of a metallic reflection on the oxygen sample at 120 GPa can be compared with the reflection on the platinum electrodes. The electrical resistance was measured using the four terminals with an applied a.c. measuring current of 1 μA. For the sample chamber with a diameter of 50 μm, we drilled a hole in an Al₂O₃ layer which was fitted into a rhenium gasket. The Al₂O₃ layer electrically insulates the metal gasket from both the sample and from the electrodes. We checked the insulation between the electrodes and the gasket during compression and all processes. We attached the diamond anvil cell and loaded the sample directly onto the mixing chamber of a ³He/⁴He dilution refrigerator. The sample temperature is expected to be close to that of the cell because the sample has negligible heat capacity and metallic thermal conductivity.

Figure 2 shows the relative resistance (R/R_{1K}) of oxygen as a function of temperature at pressures of 98, 120 and 125 GPa. As the temperature decreases, the temperature dependence shows metallic

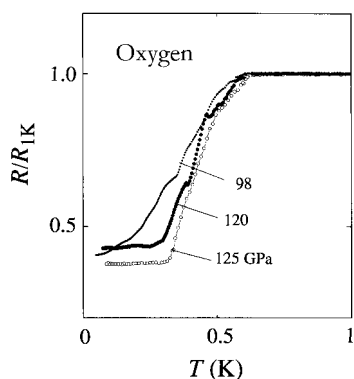


Figure 2 Relative resistance (R/R_{1K}) of oxygen as a function of temperature below 1 K at pressures of 98 (small circles), 120 (large filled circles) and 125 GPa (open circles). Drops in resistance are found at a temperature of 0.6 K. The drop is due to a loss of resistance from the superconducting transition of oxygen. The residual resistance below 0.6 K is from a part that is non-superconducting in the sample.

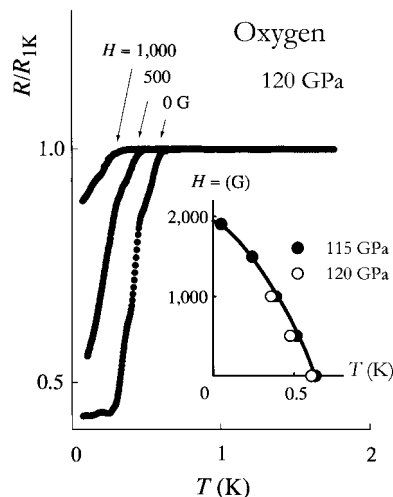


Figure 3 Magnetic field dependence of the transition at 120 GPa. Inset, critical magnetic field versus temperature diagram of oxygen at 115 GPa (filled circles) and 120 GPa (open circles). The superconductivity is suppressed and disappears under a magnetic field of 2,000 G.

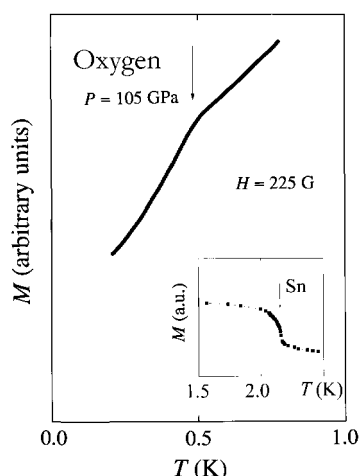


Figure 4 Detected Meissner signal from the superconducting oxygen at a magnetic field of 225 G. Superconducting signal from oxygen is detected with a kink evident at around 0.5 K. The superconducting transition of tin survives in this magnetic field and is seen at around 2 K (at 225 G) in the opposite direction to the oxygen signal, as shown in the inset.

behaviour, with a slight increase of resistance below 20 K. We observed an abrupt drop at a temperature of 0.6 K. The drop became sharper as pressure increased and a step of residual resistance can be seen. The drop is most probably attributed to the superconducting transition. Increasing the pressure causes an increase in the parts of the sample that are superconducting with almost the same superconducting transition temperature (T_c) of 0.6 K. The residual resistance below 0.6 K is from a part that is non-superconducting. This part is due to some pressure distribution in the sample. We measured the magnetic field dependence of the drop and found that the drop is suppressed and disappears under a magnetic field of 2,000 G (Fig. 3). The inset shows the critical field dependence of the transition temperature.

The Meissner measurement is indispensable to prove superconductivity. We confirmed the superconductivity by detecting the Meissner signal as follows. Sample oxygen with the same purity as the previous sample was loaded into a hole with a diameter of 60 μm drilled into a rhenium gasket. We wound several turns of superconducting pick-up coil for a SQUID magnetometer closely around the diamond and for background compensation we made a compensation coil using an equal number of turns near the diamond. The compensation coil was placed in series with the pick-up coil, and a small tin chip was inserted into the former. Tin also undergoes a superconducting transition, so we can use it to check that the chip's diamagnetic signal is in the opposite direction to that of the oxygen sample, as expected for the series coil arrangement. Figure 4 shows the Meissner effect for the superconducting oxygen at a magnetic field of 225 G. The magnetic field of 225 G completely suppressed the superconducting shield of a rhenium gasket. Rhenium under pressure is known to give a lower critical field than that at ambient pressure. Thus, the signal can be detected from oxygen that is surrounded by the rhenium gasket. We detected the transition signal from oxygen with a kink evident at around 0.5 K. The superconducting transition of tin survives in this magnetic field and is seen at around 2 K (at 225 G) in the opposite direction to the oxygen signal, as shown in the inset. The Meissner signal disappeared when we decreased the pressure down to 50 GPa.

The transition temperature of oxygen of 0.6 K is rather low compared to others in the VIb series: S and Se are at 15 K and 4.5 K, respectively^{4,5}. Previous structural experiments¹ show that solid oxygen above 96 GPa remains in the molecular state. We conclude therefore that superconductivity of oxygen appears at the molecular metallic state. □

Received 3 July 1997; accepted 16 April 1998.

1. Akahama, Y. *et al.* New high-pressure structural transition of oxygen at 96 GPa associated with metallization in a molecular solid. *Phys. Rev. Lett.* **74**, 4690–4693 (1995).
2. Desgreniers, S., Vohra, Y. K. & Ruoff, A. L. Optical response of very high density solid oxygen to 132 GPa. *J. Phys. Chem.* **94**, 1117–1122 (1990).
3. Shimizu, K. *et al.* Oxygen under high pressure; temperature dependence of electrical resistance. *Rev. High Pressure Sci. Technol.* **7**, 784–786 (1998).
4. Kometani, S. *et al.* Observation of pressure-induced superconductivity of sulfur. *J. Phys. Soc. Jpn* **66**, 2564–2565 (1997).
5. Akahama, Y. *et al.* Pressure-induced superconductivity and phase transition in selenium and tellurium. *Solid State Commun.* **84**, 803–806 (1992).

Acknowledgements. We thank T. C. Kobayashi for helpful discussion. This work was partly supported by a Grant-in-Aid for Scientific Research (A) from the Ministry of Education, Science, Sports and Culture, Japan.

Correspondence and requests for materials should be addressed to K.S. (e-mail: kshimizu@mp.es.osaka-u.ac.jp).

Glass fibres of pure and erbium- or neodymium-doped yttria–alumina compositions

J. K. Richard Weber, John J. Felten, Benjamin Cho & Paul C. Nordine

Containerless Research, Inc., 906 University Place, Evanston, Illinois 60201, USA

Optical fibres doped with lanthanide or transition-metal elements can serve as in-line lasers and amplifiers for fibre-optic telecommunications systems. In general, most such fibre lasers use conventional silica-glass fibres doped with erbium or neodymium. But silicon dioxide absorbs strongly in the infrared for wavelengths of greater than 4 μm or so, limiting the infrared range over which such lasers can operate. Some other oxide materials do not absorb significantly until longer wavelengths—the absorption coefficient of crystalline silica at 4 μm is equal to that of yttrium oxide at 7.1 μm and of sapphire (a form of alumina) at 5.1 μm , for example¹. Glass fibres made from these materials would therefore expand the range of fibre lasers into the mid-infrared. But molten oxides that do not contain silica typically have a viscosity too low to support fibre-pulling processes. Here we demonstrate that containerless processing, in which a molten sample is levitated by a flow of inert gas, permits sufficient undercooling of molten yttrium aluminium garnet (YAG:Y₃Al₅O₁₂) to access a viscosity range conducive to fibre-pulling. The process is particularly effective if the molten material of stoichiometric YAG composition is doped with Nd₂O₃ in place of Y₂O₃, or with excess Al₂O₃; and it should also work with other dopants, because molten oxides are good solvents. Fibres could be drawn from a melt doped with Er₂O₃ in the presence of excess Al₂O₃. These

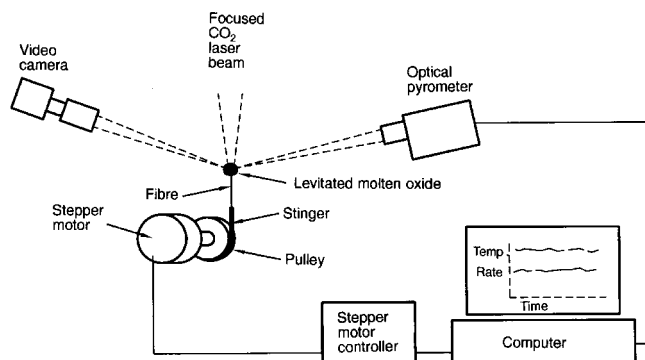


Figure 1 Diagram of the apparatus used to make fibres from levitated drops of undercooled molten oxides.

fibres have the potential to extend the operating range of oxide glass-fibre lasers.

Liquids that form glasses are characterized as 'strong' or 'fragile' glass-formers²⁻⁴. Strong liquids such as molten SiO₂ show an Arrhenian temperature dependence of viscosity over a wide temperature range. Fragile liquids depart from this temperature dependence and have viscosities at lower temperatures that are much larger than predicted by extrapolation of viscosity measurements obtained at high temperatures. It has been clear for some time that molten YAG is a fragile liquid. Fratello and Brandle⁵ report viscosities near to and above the melting point that are of the order of 0.05 Pa s and only weakly dependent on temperature. The liquid can be deeply undercooled to form a glass under containerless melt-processing conditions⁶. The glass transition temperature, T_g , of YAG-composition glass was estimated to be 1,273 K from thermal analysis measurements (P. F. McMillan, personal communication). The viscosity of the liquid at T_g was taken to be 10¹² Pa s. However, extrapolation of the high-temperature viscosity data⁵ to 1,273 K yields values of only a few pascal seconds.

We have used containerless liquid-phase processing to achieve deep undercooling of molten oxides to temperatures where the viscosity is sufficient for fibre-pulling operations. The containerless conditions eliminate heterogeneous nucleation by containers, and allow deep undercooling to occur⁷. The experimental apparatus is illustrated in Fig. 1. Specimens were levitated in a flow of argon gas using the conical nozzle levitation technique⁷. Levitated material was heated and melted with a continuous-wave CO₂ laser heating

beam. The sample temperature was measured at a rate of 30 Hz with an automatic optical pyrometer. The samples, of ~3 mm diameter, were synthesized from high-purity oxide powders by laser hearth melting⁸. The compositions investigated are presented in Table 1.

The levitated samples were completely melted and the heating laser beam was then blocked, resulting in cooling at a rate of ~250 K s⁻¹. The fibres were pulled from the levitated liquid drop by rapidly introducing and withdrawing a 100-μm-diameter tungsten wire 'stinger' at a preselected temperature. The fibre-pulling operations were performed under computer control using the pyrometer output to trigger the stinging operation. The direction of stinger motion was reversed a few milliseconds after it entered the liquid drop: longer residence times led to crystallization of the melt. As fibre-pulling progressed, the stinger and fibre wrapped around the pulley attached to the motor spindle (Fig. 1). Fibre production was stopped when the fibre-pulling force brought the liquid drop into contact with the levitation nozzle structure, and nucleation of crystals and crystallization of the melt occurred.

Glass fibres up to 0.5 m long were pulled at rates of 1–1.5 m s⁻¹ before crystallization of the melt terminated the process. The melt temperature decreased by ~75 K between fibre stinging and crystallization of the melt. Fibres were obtained only if stinging was initiated when the melt temperature was in the range 1,600–1,660 K. At higher temperatures, the stinger pulled out of the melt without forming a fibre; at lower temperatures, the drop was pushed away by the stinger.

The fibres had a homogeneous appearance with smooth surfaces;

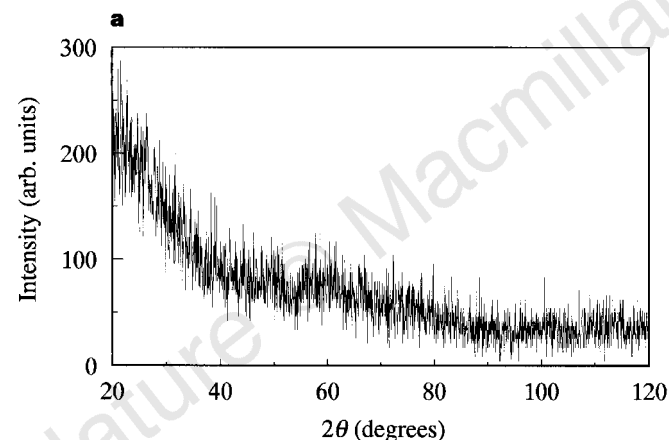


Figure 2 Properties of fibres produced in our experiments. **a**, X-ray diffraction pattern obtained from a sample of finely ground fibre containing 1% neodymium oxide substituted for yttrium oxide. **b**, Photograph of glass fibres pulled from undercooled molten YAG-composition materials. The fibres were pulled from levitated drops at 1,600–1,660 K, ~600 K below the melting point of crystalline YAG. The fibres are ~20 μm in diameter.

Table 1 Composition of oxide materials from glass-fibre-pulling experiments

Al ₂ O ₃ (mol%)	Y ₂ O ₃ (mol%)	Comments
62.5	37.5	From mixed powders and from crushed single-crystal YAG
63.5	36.5	From mixed oxide powders
62.5	36.5	1 mol% Nd ₂ O ₃
62.5	36.5	1 mol% Er ₂ O ₃
63.5	35.5	1 mol% Er ₂ O ₃

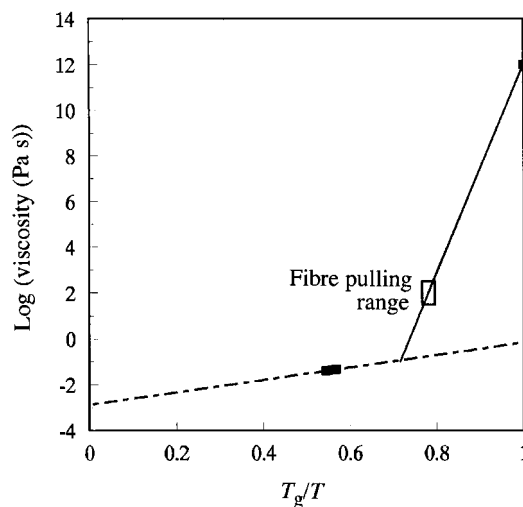


Figure 3 Data on the temperature dependence of the viscosity of molten YAG. Shown is a plot of the logarithm of viscosity versus the ratio of its glass-transition temperature (T_g) to its absolute temperature (T). The plot is based on a T_g value of 1,273 K. The temperature range from 1,600 to 1,660 K and the viscosity range from 10¹⁵ to 10^{2.5} Pa s required for fibre-pulling are indicated by the boxed region. The solid, steep line connects the viscosity values at T_g (black square) with the fibre-pulling range. The dot-dashed, lower line is the least-squares fit to the viscosity data of Fratello and Brandle⁵ (black rectangle), extrapolated over the temperature range of the plot. The extrapolation shows that the viscosity of the undercooled liquid would be only ~1 Pa s at $T = T_g$ and slightly smaller at the fibre-pulling temperatures.

they were transparent and highly flexible, and showed a glassy fracture surface. The X-ray diffraction pattern of crushed fibre material (Fig. 2a) confirmed that it was glassy. This result is consistent with the report by Massiot *et al.*⁶ that glassy YAG can be formed from the undercooled liquid. A photograph of several glass fibres is shown in Fig. 2b.

The initial experiments were performed on mixtures of the approximate $Y_3Al_5O_{12}$ (YAG) composition prepared from pure Al_2O_3 and Y_2O_3 powders. Some of the specimens readily formed long fibres whereas others did not, even though the melts were undercooled into the fibre-pulling temperature regime. In some cases only very short fibres of <1 mm in length were formed at the end of the stinger.

Later experiments showed that the variations in the fibre-pulling behaviour resulted from small differences in the compositions of the synthetic specimens. Fibres were rarely obtained if the melt was formed from stoichiometric single-crystal YAG, but long fibres were reproducibly pulled from compositions with an excess of 1 mol% Al_2O_3 . Fibres were readily obtained from molten YAG compositions containing 1 mol Nd_2O_3 substituted for Y_2O_3 . However, if Er_2O_3 was substituted for Y_2O_3 the behaviour was similar to that of stoichiometric YAG and fibres were rarely obtained. Addition of an excess of 1 mol% Al_2O_3 to the Er-containing material resulted in reproducible pulling of long glass fibres.

The initial sections of the fibres contained small nodules with smoothly varying diameters up to about four times the nominal fibre diameter and lengths that were a few times the nodule diameter. The central region of the long glass fibres was of uniform diameter, typically in the range 5–30 μm . The terminal sections of the fibres showed occasional nodular defects that were smaller, and appeared more abruptly than those in the initial section.

A further difference between the pure YAG and doped compositions was found in cooling experiments in which fibre-pulling was not conducted. Pure YAG compositions, and those with Er_2O_3 substituted for Y_2O_3 always crystallized when cooled below 1,400 K, but the compositions with excess Al_2O_3 or Nd_2O_3 substituted for Y_2O_3 formed a glass. In addition, the crystallization behaviour of pure YAG was sensitive to preheating conditions, such that either crystalline YAG or a mixture of $YAlO_3$ and Al_2O_3 could be formed.

In Fig. 3 we show data on the viscosity of molten YAG, as a semi-log plot of viscosity versus the ratio of the glass-transition temperature to the absolute temperature. The range of viscosities shown for fibre-pulling was taken to be that typical for glass-fibre manufacture, that is, $10^{1.5}$ to $10^{2.5}$ Pa s (ref. 9). This is the viscosity of the bulk liquid whose temperature was measured, and we recognize that further cooling of the liquid—and a further increase in viscosity—occurs in the meniscus leading to the fibres that form. This result is obtained for all of the compositions investigated. The sources of other data shown in Fig. 3 are given in the figure legend.

The lines in Fig. 3 confirm that molten YAG is a fragile glass-former. The viscosity versus T_g/T behaviour is similar to that of *o*-terphenyl, which is known to be an extremely fragile liquid¹⁰. We remark that, unlike many other glasses, both the present results and McMillan's measurement show that the value of T_g is well below two-thirds of the melting point of crystalline YAG.

There are at least two phenomena that may affect the fibre-pulling behaviour observed for pure YAG and doped compositions. The first is the greater sensitivity of undercooling behaviour in general for the pure composition. It is well known that YAG crystal growth can encounter difficulties due to crystallization of other phases^{11,12}. Evidently, undercooled pure YAG melts are sensitive to small compositional or other differences associated with preheating, and this sensitivity may also influence the glass-fibre-pulling process.

The second phenomenon that may affect fibre-pulling was reported by Aasland and McMillan¹³. They found that a mixture of two glasses forms on rapid cooling of Al_2O_3 – Y_2O_3 melts contain-

ing 24–32 mol% Y_2O_3 . The difficulty in making pure YAG fibres and the ease with which doped fibres were formed may result from different phase-transition kinetics for the pure and doped materials. In addition, the abrupt change in fibre-stinging behaviour at 1,600 K has characteristics of a phase transition from a smaller- to a much larger-viscosity liquid.

Glass-fibre synthesis should be possible for a wide range of oxide materials. We have also made glass fibres from alumina–silica compositions with Al_2O_3 : SiO_2 up to 72:28, Mg_2SiO_4 (forsterite), $CaAl_2O_4$, $LaAlO_3$ and others. Although viscosity data for non-silicate oxides are lacking, it is apparent from glass-forming results that many, if not all, of these liquids are highly fragile. The method of containerless undercooling to access the higher viscosities required for glass-fibre-pulling should thus be applicable to the manufacture of fibres of a large number of new oxide compositions. □

Received 23 January; accepted 7 April 1998.

1. Bass, M. (ed.) *Handbook of Optics* Vol. 2, *Devices, Measurements and Properties* 2nd edn (McGraw-Hill, New York, 1995).
2. Angell, C. A. in *Relaxations in Complex Systems* (eds Ngai, K. L. & Wright, G. B.) 3–11 (National Technical Information Service, US Dept of Commerce, Springfield, VA, 1985).
3. Angell, C. A. Formation of glasses from liquids and biopolymers. *Science* **267**, 1924–1935 (1995).
4. Debenedetti, P. G. *Metastable Liquids* (Princeton Univ. Press, 1997).
5. Fratello, V. J. & Brandle, C. D. Physical properties of a $Y_3Al_5O_{12}$ melt. *J. Cryst. Growth* **128**, 1006–1010 (1993).
6. Massiot, D., Taulelle, F. & Coutures, J. P. Structure diagnostic of high temperature liquid phases by ²⁷Al NMR. *Colloque Phys. C 5*, 425–431 (1990).
7. Weber, J. K. R. & Nordine, P. C. Containerless liquid-phase processing at high temperatures. *Microgravity Sci. Technol.* **VII**, 279–282 (1995).
8. Weber, J. K. R., Felten, J. J. & Nordine, P. C. New method for high purity ceramic synthesis. *Rev. Sci. Instrum.* **67**, 522–524 (1996).
9. Loewenstein, K. L. *The Manufacturing Technology of Continuous Glass Fibres* (Elsevier, Amsterdam, 1973).
10. Laughlin, W. T. & Uhlmann, D. R. Viscous flow in simple organic liquids. *J. Phys. Chem.* **76**, 2317–2325 (1972).
11. Abell, J. S., Harris, I. R., Cockayne, B. & Lent, B. An investigation of phase stability in the Y_2O_3 – Al_2O_3 system. *J. Mater. Sci.* **9**, 527–537 (1974).
12. Caslavsky, J. L. & Viechnicki, D. J. Melting behavior and metastability of yttrium aluminum garnet (YAG) and $YAlO_3$ determined by optical differential thermal analysis. *J. Mater. Sci.* **15**, 1709–1718 (1980).
13. Aasland, S. & McMillan, P. F. Density-driven liquid-liquid phase separation in the system Al_2O_3 – Y_2O_3 . *Nature* **369**, 633–639 (1994).

Acknowledgements. We thank P. F. McMillan and C. A. Angell for discussions. This work was supported by NASA. B. Cho was partially supported by the State of Illinois, USA.

Correspondence and requests for materials should be addressed to J.K.R.W. (e-mail: weber@containerless.com).

C₃₆, a new carbon solid

C. Piskoti^{*†}, J. Yarger^{‡†} & A. Zettl^{*†}

^{*} Department of Physics, University of California at Berkeley, 366 LeConte Hall, Berkeley, California 94720, USA

[†] Materials Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA

[‡] Department of Chemistry, University of California at Berkeley, 420 Latimer Hall, Berkeley, California 94720, USA

Under appropriate non-equilibrium growth conditions, carbon atoms form relatively stable hollow clusters of well-defined mass number¹, collectively known as fullerenes. The mass production, purification and condensation of such clusters into a molecular solid is generally essential to full experimental characterization: the initial discovery² of C_{60} , for example, had to await a bulk synthesis method³ six years later before detailed characterization of the molecule was possible. Gas-phase experiments^{1,4,5} have indicated the existence of a wide range of fullerene clusters, but beyond C_{60} only a few pure fullerene solids have been obtained⁶, most notably C_{70} . Low-mass fullerenes are of particular interest because their high curvature and increased strain energy owing to adjacent pentagonal rings could lead to solids with unusual

intermolecular bonding and electronic properties. Here we report the synthesis of the solid form of C_{36} by the arc-discharge method². We have developed purification methods that separate C_{36} from amorphous carbon and other fullerenes, to yield saturated solutions, thin films and polycrystalline powders of the pure solid form. Solid-state NMR measurements suggest that the molecule has D_{6h} symmetry, and electron-diffraction patterns are consistent with a tightly bound molecular solid with an intermolecular spacing of 6.68 Å. We observe large increases in the electrical conductivity of the solid on doping with alkali metals.

To determine the optimum parameters for C_{36} production, we performed a series of experiments in a helium-environment arc-discharge chamber originally designed for C_{60} production. An arc was started between two 1/4-inch-diameter graphite electrodes, using a d.c. current of 100 A while maintaining a 1-mm gap between electrodes. A removable metal substrate, designed for subsequent direct insertion into a mass spectrometer, was placed 10 cm from the discharge region on the wall of the water-cooled chamber. Arcing was maintained for several minutes until a uniform carbon film of approximate thickness 10 µm coated the substrate. A series of experimental runs was carried out at different fixed static helium pressures between 50 and 1,500 torr. Each of the substrates was then

loaded into a Micromass laser desorption/ionization time-of-flight mass spectrometer.

Figure 1a shows the mass spectrum of one of these films grown in 400 torr helium. The dominant peaks in the spectrum are at 720 and 432 atomic mass units (a.m.u.). The former peak corresponds to 60 carbon atoms, while the latter corresponds to 36 carbon atoms. The C_{60} and C_{36} peaks are of comparable magnitude and these molecular species appear to be the most prominent in the sample. Lesser peaks are also observed, for example at 840 a.m.u. (C_{70}). The synthesis of C_{36} is very sensitive to experimental conditions, notably the helium pressure: runs in this series at pressures significantly different from 400 torr failed to produce prominent peaks below 720 a.m.u. in the mass spectrum.

To produce bulk amounts of C_{36} suitable for purification, arcing runs in 400 torr helium were repeated and the resulting 'soot' was collected from the synthesis chamber walls. The soot was initially washed with toluene in a standard Soxhlet extractor which removed C_{60} , C_{70} and trace amounts of even higher-order fullerenes. Mass spectrometry on the toluene-soluble extract showed the expected fullerene peaks but no peak at 432 a.m.u., indicating that C_{36} is not soluble in toluene.

Two methods were subsequently used to separate C_{36} from the remaining toluene-insoluble material. In the first, sublimation-purification was used to produce thin films of C_{36} . Here the toluene insoluble material was dried in vacuum at 150 °C for two hours and placed in the tungsten boat of a thermal evaporator. No precautionary steps were taken to protect the material from ambient moisture. The evaporator boat was heated under a pressure of less than 1 µtorr, ultimately reaching a temperature in excess of 1,500 °C. A metal substrate was suspended over the boat at a height of about 5 cm. In about 20 minutes, films of the order of 100 nm thick could be grown on the substrate. Notably, the black films were highly resistant to scratching, unlike amorphous carbon or C_{60} films.

Figure 1b shows the time-of-flight mass spectrum for the material of a thermally evaporated film. The film contains essentially a single molecular component of mass 438 a.m.u., 6 a.m.u. higher than the mass of C_{36} . Almost certainly the discrepancy is due to the partial hydrogenation of the C_{36} clusters through reaction with adsorbed moisture at the elevated temperature of evaporation. As expected, C_{36} is a fairly reactive material, more so than C_{60} or C_{70} .

To establish a purification method based on 'wet chemistry' rather than thermal evaporation, we tested sublimation-generated films prepared as above in a variety of organic solvents. The films were found to be insoluble in toluene and benzene, but soluble in pyridine and carbon disulphide (CS_2). However, unlike C_{60} and C_{70} , C_{36} is very slow to dissolve in either of these solvents and extensive heating and exposure to ultrasound are required to obtain a saturated solution. This is consistent with the strong, almost covalent bonding expected⁷ between C_{36} molecules: C_{36} is not simply a van der Waals solid as is C_{60} .

The wet-chemistry purification method was used to produce bulk quantities of refined C_{36} powder. The crude arc-chamber soot was first Soxhlet-extracted with toluene to remove higher-order fullerenes, followed by Soxhlet extraction with pyridine. The pyridine-soluble extract forms a yellow-brown solution, as opposed to the toluene-soluble extract which is red-brown in colour. Overall, the crude soot was found to contain ~10% toluene-soluble material (C_{60} and C_{70}) and about 1–2% pyridine-soluble material (C_{36}). After evaporating the pyridine from the pyridine-soluble fraction, a black solid was obtained. This solid differs from that of the previously reported $C_{36}H_{12}$ bowl-shaped molecule in both colour and solubility⁸.

To investigate the structure of the individual C_{36} molecules and the C_{36} molecular solid, we used solid-state ^{13}C NMR, bulk electron diffraction, mid-infrared transmission, and solid-state transport studies. It has been shown that for a 36-atom carbon cage with

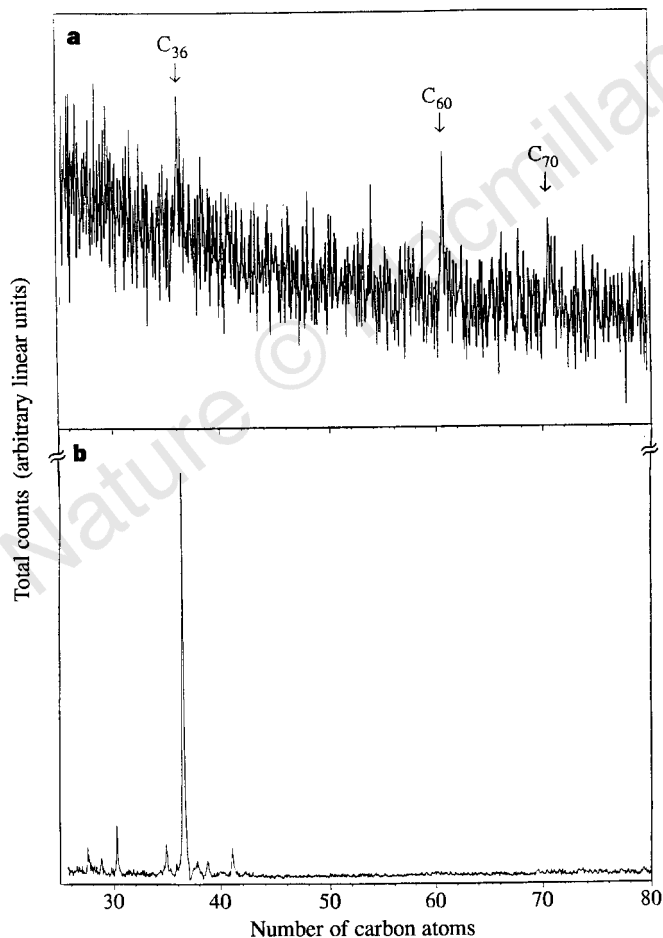


Figure 1 Mass spectra of films prepared in this study. **a**, Time-of-flight mass spectrum of a sample of crude fullerene soot produced at a helium pressure of 400 torr. The data show the presence of a peak at the mass of C_{36} (432 a.m.u.) with approximately the same intensity as the peak due to C_{60} (720 a.m.u.). **b**, Time-of-flight mass spectrum of material from a film evaporated from crude soot which had been depleted of C_{60} and C_{70} by toluene extraction. The peak at 438 a.m.u. is due to C_{36} which has been partially hydrogenated (six hydrogens per C_{36} molecule) during the high-temperature evaporation process.

hexagonal and pentagonal faces, 15 different isomers are possible⁹. Recent calculations indicate that of those isomers the lowest in energy are the structures with D_{6h} and D_{2d} symmetry¹⁰. These two molecules can be distinguished via NMR spectroscopy. The D_{6h} structure is expected to show three distinct ^{13}C NMR resonances of equal intensity, whereas the D_{2d} structure is expected to have five resonances, four with equal intensity and the fifth with half the intensity^{9,11}.

Figure 2 shows the experimental ^{13}C NMR spectrum for C_{36} powder, obtained using a Chemagnetics 500 MHz instrument with magic-angle spinning. The experimental spectrum contains two prominent peaks, one at 146.1 p.p.m. (relative to tetramethylsilane) and another (with approximately one-half the intensity) at 135.7 p.p.m. The inset to Fig. 2 shows the predicted molecular NMR spectra for the isolated D_{6h} and D_{2d} isomers (along with schematic structure drawings). The experimental spectrum appears inconsistent with predictions for the D_{2d} isomer. On the other hand, taking into account experimental broadening of the peaks, one would expect for the D_{6h} isomer two peaks, one near 135 p.p.m. and another 'double intensity' peak at higher p.p.m. arising from the two higher nearly degenerate resonances. This is precisely what is observed experimentally. The smaller experimentally observed shift of the 'double intensity' peak (at 146 p.p.m. versus the predicted 158 p.p.m.) is accounted for by additional shielding of these reactive carbon atom sites by neighbouring molecules in the solid (this shielding is not considered in the simple molecular calculations). We thus identify our C_{36} cage molecule as having D_{6h} symmetry.

To investigate the crystal structure of solid C_{36} , we performed electron diffraction experiments. A small amount of material obtained from the pyridine extraction was ground up, dispersed on a holey carbon grid, and inserted into a JEOL 200X transmission electron microscope (TEM). The material was observed to be polycrystalline with a crystallite grain size of ~ 100 nm. Using a field limiting aperture, the diffraction pattern of selected crystallites was recorded.

In Fig. 3 we show a TEM diffraction pattern for a C_{36} crystallite. The hexagonal diffraction pattern suggests a close-packing arrangement perpendicular to the zone axis. The pattern is reminiscent of diffraction patterns observed for C_{60} and C_{70} . However, the d -spacing measured from this pattern for the first-order diffracted

spots is 6.68 Å, significantly less than the (100) d -spacing of 8.7 Å reported³ for C_{60} . Unfortunately, because the C_{36} crystallites are platelets with high aspect ratios, this was the only zone axis along which the crystallites were thin enough to allow useful TEM imaging, and thus determination of the detailed C_{36} crystal structure was not possible. Using the C_{36} molecular diameter^{7,10,11} of approximately 5 Å (carbon centre to carbon centre distance) and assuming that C_{36} crystallizes in close-packed planes, the diffraction results imply an intermolecular bond length of 1.7 Å, significantly smaller than the van der Waals distance of 3.4 Å in graphite. This is suggestive of strong crystal bonding.

Like higher-order fullerene solids, C_{36} appears to suffer some TEM-induced damage at an electron beam energy of 200 keV. Under continuous TEM observation, the sharp crystalline diffraction patterns were found to deteriorate for extended irradiation times. Additional studies were performed to investigate the long-term stability of solid C_{36} subject (only) to high ambient temperatures. C_{36} powder was heated in vacuum to 1,350 °C for 48 hours and then characterized by TEM imaging. Over 50% of the diffraction patterns obtained for the heat-treated material were graphitic, indicating that a large amount of the material had converted to the energetically more favourable graphite.

Specimens of solid C_{36} were further characterized by mid-infrared transmission and d.c. electrical resistivity measurements. Figure 4 shows the room temperature mid-infrared transmission spectrum for C_{36} powder dispersed in KBr. Several prominent dips in the transmission are observed. In the simplest interpretation, these dips directly represent absorption bands for the C_{36} solid. The observed absorption resonances are in the fullerene range but, as expected, they do not precisely match those of previously known fullerenes. No additional absorption peaks for C_{36} were observed in the frequency range 2,000–5,000 cm^{-1} .

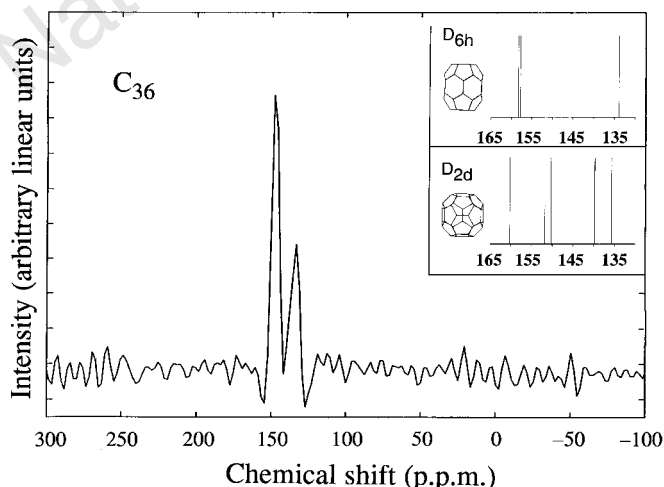


Figure 2 Predicted and observed NMR spectra of C_{36} . Main figure, ^{13}C NMR spectrum of C_{36} powder showing peaks at 146.1 and 135.7 p.p.m. with a 2:1 relative intensity ratio. Inset, the two lowest-energy isomers of the C_{36} molecule (with D_{2d} and D_{6h} symmetry) with the corresponding predicted NMR spectra¹². Taking into account broadening and shielding for the solid, the experimental data suggest D_{6h} symmetry for the C_{36} molecule.



Figure 3 Electron-diffraction pattern of C_{36} crystallite. The pattern is hexagonal with a calculated d -spacing of 6.68 Å.

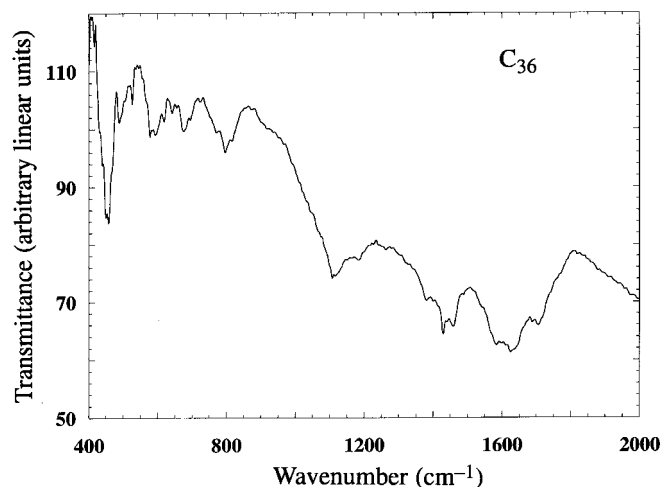


Figure 4 Infrared transmission spectrum of C_{36} powder in KBr.

For d.c. measurements, four-probe contacts were made to polycrystalline specimens with conductive silver paint. At room temperature, the pure samples were insulating with a resistivity in excess of $10^7 \Omega \text{ cm}$. In fullerenes such as C_{60} , it is known that intercalation with alkali metals or alkaline earths drastically improves the electrical conductivity of the solid^{1,12}. We have investigated this effect for C_{36} . Intercalation of the C_{36} solid with alkali metals was carried out using both sodium and potassium, by sealing the sample at one end of an evacuated tube with the alkali metal at the other end. The sample was then heated to $\sim 250^\circ \text{C}$ while the metal was heated to $\sim 175^\circ \text{C}$. In both cases, the resistivity dropped by more than four orders of magnitude, to $\sim 10^3 \Omega \text{ cm}$ at room temperature. The resistivity ρ as a function of temperature T for samples intercalated with either sodium or potassium was not metallic, but obeyed the functional form $\rho \sim \exp[T^{-\alpha}]$, behaviour typical of variable range hopping. At present it is not clear where in the structure, or how uniformly, the alkali metals intercalate.

If C_{36} could be made conducting with a sufficient density of states at the Fermi level—either by doping or by structural rearrangement (for example, induced by pressure)—one might expect, among other things, high-temperature superconductivity¹⁰ with a transition temperature significantly exceeding that of intercalated C_{60} . These possibilities, along with details of the structure, electronic properties, and mechanical response of C_{36} , are at present being investigated. $\square$

Received 6 April; accepted 21 April 1998.

- Billups, W. E. & Ciufolini, M. A. (eds) *Buckminsterfullerenes* (VCH, New York, 1993).
- Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. & Smalley, R. E. C_{60} : buckminsterfullerene. *Nature* **318**, 162–163 (1985).
- Kratschmer, W., Lamb, L. D., Fostiropoulos, K. & Huffman, D. R. Solid C_{60} : a new form of carbon. *Nature* **347**, 354–358 (1990).
- Rohlfing, C. & Kaldor, J. Production and characterization of supersonic carbon cluster beams. *Chem. Phys.* **81**, 3322–3330 (1984).
- O'Brien, S. C., Heath, J. R., Curl, R. F. & Smalley, R. E. Photophysics of buckminsterfullerene and other carbon cluster ions. *J. Chem. Phys.* **88**, 220–230 (1988).
- Diederich, F. & Whetten, R. L. Beyond C_{60} : the higher fullerenes. *Acc. Chem. Res.* **25**, 119–126 (1992).
- Côté, M., Grossman, J. C., Louie, S. G. & Cohen, M. L. Electronic and structural properties of molecular C_{36} . *Bull. Am. Phys. Soc.* **42**, 270 (1997).
- Scott, L. T., Bratcher, M. S. & Hagen, S. Synthesis and characterization of a C_{36} subunit. *J. Am. Chem. Soc.* **118**, 8743–8744 (1996).
- Fowler, P. W. & Manolopoulos, D. E. *An Atlas of Fullerenes* (Clarendon, Oxford, 1995).
- Grossman, J. C., Côté, M., Louie, S. G. & Cohen, M. L. Prediction of superconductivity in solid C_{36} . *Bull. Am. Phys. Soc.* **42**, 1576 (1997).
- Grossman, J. C., Côté, M., Louie, S. G. & Cohen, M. L. Electronic and structural properties of molecular C_{36} . *Chem. Phys. Lett.* **284**, 344–349 (1998).
- Hebard, A. F. et al. Superconductivity at 18 K in potassium doped C_{60} . *Nature* **350**, 600–601 (1991).

Acknowledgements. The study of C_{36} arose out of a joint experimental/theoretical collaboration with M. L. Cohen, M. Côté, J. C. Grossman and S. G. Louie, all of whom we thank for discussions and interactions. We also thank J. Burward-Hoy for contributing to the initial stages of this work, T. Wagberg and M. C. Martin for providing the infrared spectrum, A. Pines for discussions and use of his NMR spectrometer, and J. O'Leary for performing the mass spectroscopy runs on the Micromass, Inc. equipment. This work was supported in part by the US Department of Energy, the US Office of Naval Research, and the US NSF.

Correspondence and requests for materials should be addressed to (e-mail: azettl@physics.berkeley.edu).

Influence of iron availability on nutrient consumption ratio of diatoms in oceanic waters

Shigenobu Takeda

Biology Department, Central Research Institute of Electric Power Industry, Abiko, Chiba 270-11, Japan

The major nutrients (nitrate, phosphate and silicate) needed for phytoplankton growth are abundant in the surface waters of the subarctic Pacific, equatorial Pacific and Southern oceans, but this growth is limited by the availability of iron^{1–5}. Under iron-deficient conditions, phytoplankton exhibit reduced uptake of nitrate⁶ and lower cellular levels of carbon, nitrogen and phosphorus⁷. Here I describe seawater and culture experiments which show that iron limitation can also affect the ratio of consumed silicate to nitrate and phosphate. In iron-limited waters from all three of the aforementioned environments, addition of iron to phytoplankton assemblages in incubation bottles halved the silicate:nitrate and silicate:phosphate consumption ratios, in spite of the preferential growth of diatoms (silica-shelled phytoplankton). The nutrient consumption ratios of the phytoplankton assemblage from the Southern Ocean were similar to those of an iron-deficient laboratory culture of Antarctic diatoms, which exhibit increased cellular silicon or decreased cellular nitrogen and phosphorus in response to iron limitation. Iron limitation therefore increases the export of biogenic silicon, relative to nitrogen and phosphorus, from the surface to deeper waters. These findings suggest how the sedimentary records of carbon and silicon deposition in the glacial Southern Ocean⁸ can be consistent with the idea that changes in productivity, and thus in drawdown of atmospheric CO_2 , during the last glaciation were stimulated by changes in iron inputs from atmospheric dust.

Iron-enrichment bottle incubation experiments were performed with resident phytoplankton assemblages to determine the effect of iron on major-nutrient uptake in three upwelling areas of the open ocean where iron input from atmospheric dust is known to be low⁹; the Southern Ocean, the equatorial Pacific, and the subarctic North Pacific.

In the Southern Ocean, iron addition promoted chlorophyll *a* increase and nitrate decrease compared to the control, indicating iron limitation of phytoplankton growth and nutrient utilization (Fig. 1a, b). However, significant decrease of silicate was observed both in the iron-enriched bottle and in the control (Fig. 1c): the elemental ratio of nutrients consumed by the phytoplankton assemblage in the control bottle showed silicate:nitrate and silicate:phosphate ratios two times higher than that in the iron-enriched bottle (Table 1). Yet microscopic examinations showed substantial growth of large-size diatoms (*Chaetoceros* spp. and *Nitzschia* spp.) in the iron-enriched bottle. The phytoplankton assemblage in the control was composed mainly of the same species as in the iron-enriched bottle. Thus, the iron nutritional status of the diatoms appears to affect silicate utilization physiologically.

The iron-induced change in nutrient consumption ratio was also observed in a phytoplankton assemblage from the Southern Ocean incubated under low light intensity (Table 1). Low light tended to decrease the phytoplankton growth rate and nutrient utilization, while the chlorophyll *a* content per cell increased. Even under such light-limited conditions, iron enrichment halved the consumption ratio of silicate to nitrate and phosphate (Table 1). Thus, the iron nutritional status appears to affect the consumption ratio of major nutrients in the whole euphotic zone.

Essentially the same results were observed in waters taken from

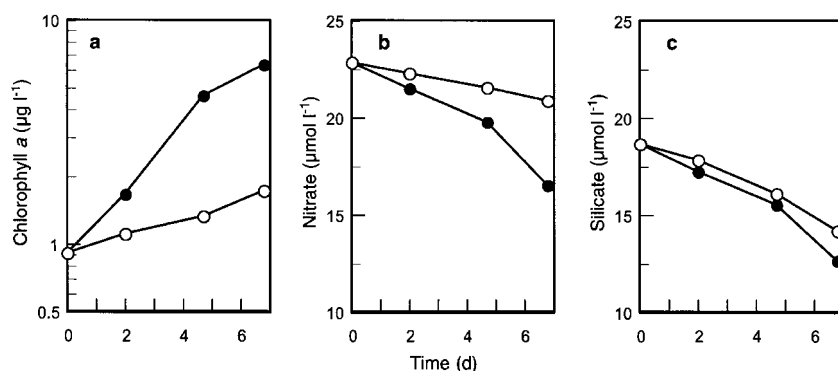


Figure 1 Results of *in vitro* iron-enrichment experiment in the Southern Ocean (64.2° S, 140.7° E; Jan. 1995). Shown are concentrations of chlorophyll *a* (**a**), nitrate (**b**) and silicate (**c**) versus time (days since the start of the experiment). Open circles, control (no iron addition); filled circles, bottle with 1 nmol l⁻¹ iron addition.

Table 1 *In vitro* iron-enrichment experiments in three oceanic waters

Treatment	Fe ⁺ (nmol l ⁻¹)	Growth rate† (d ⁻¹)	NO ₃ ⁻ used‡ (μmol l ⁻¹)	Nutrient consumption ratio‡		
				Si(OH) ₄ :NO ₃ ⁻ (mol mol ⁻¹)	Si(OH) ₄ :PO ₄ ³⁻ (mol mol ⁻¹)	NO ₃ :PO ₄ ³⁻ (mol mol ⁻¹)
Southern Ocean (64.2° S, 140.7° E; Jan. 1995)						
Control	0.16	0.13	2.0	2.3	26	12
Fe added	1.2	0.43	6.3	0.95	14	14
Low light§	0.20	0.25	0.8	2.4	20	8.1
Low light§ and Fe added	1.3	0.40	3.2	1.0	14	14
Subarctic North Pacific (45° N, 165° E; Oct. 1993)						
Control	0.22	0.14	1.2 ± 0.6	2.6 ± 0.5	31 ± 7	12 ± 0.5
Fe added	1.1	0.49	2.5 ± 0.2	1.2 ± 0.2	14 ± 6	12 ± 3
Equatorial Pacific (0°, 159° W; Nov. 1993)						
Control	<0.05	0.06	0.3 ± 0.1	1.3 ± 0.2	30 ± 17	23 ± 9
Fe added	0.46	0.48	1.9 ± 0.4	0.45 ± 0.01	6.5 ± 0.4	14 ± 1
Fe added	0.72	0.79	2.7 ± 0.2	0.45 ± 0.03	7.6 ± 1.3	17 ± 2

* Initial concentration of total dissolvable iron in incubation bottles.

† Phytoplankton growth rate based on chlorophyll *a* concentrations.

‡ Means ± s.d. of duplicated incubation bottles except for the Southern Ocean experiments (no replicates).

§ Light intensity was reduced to 2.6% of the light level at the sea surface.

the subarctic North Pacific and the equatorial Pacific: iron additions stimulated chlorophyll *a* production and nutrient consumption, but not significantly with respect to silicate (Table 1). As a result, the silicate:nitrate and silicate:phosphate consumption ratios in the controls were twice as high as those in the iron-enriched bottles (Table 1). In contrast, nitrate:phosphate consumption ratios remained constant with and without iron enrichment. These results indicate that, in all three ocean upwelling areas, iron limitation leads to preferential drawdown of silicate from the surface mixed layer compared to nitrate and phosphate. Thus, iron should be considered not only as a factor directly limiting phytoplankton growth, nitrate utilization and biomass, but also as a factor indirectly controlling the biogeochemical cycling of nutrients, particularly silicate.

Similar nutrient consumption ratios were obtained from two high-latitude waters (Table 1). In the equatorial Pacific, however, the values are nearly a half of those observed in the high-latitude waters; the reason for this is not clear. The degree of silicification may decrease with increasing temperature in some diatom species¹⁰. The low ambient silicate concentration (2.1 µmol l⁻¹) in the equatorial Pacific surface water may also contribute to the lower silicate consumption ratio, because silicate limitation significantly affects the diatom Si:N ratio¹¹. Though the present data are restricted to oceanic upwelling regions, similar effects of iron availability on nutrient utilization were recently observed in the Californian coastal upwelling region¹². I hypothesize that iron control of silicate utilization can be extrapolated to other iron-limited environments.

Decreases in silicate:nitrate and silicate:phosphate consumption ratios of phytoplankton assemblages can arise from a shift in phytoplankton species composition—from algae with siliceous structures to algae that do not require silicon. However, such a species change is not responsible for the present results, because iron enrichment enhances the growth of diatoms in most cases. A shift in the main nitrogen source from ammonium to nitrate with iron enrichment may partly be responsible for the change in the silicate:nitrate ratio, although it can not explain the change in the silicate:phosphate ratio. Phytoplankton growing with nitrate require more iron for the synthesis of nitrate and nitrite reductase than those growing with ammonium¹³. Thus restricted nitrate uptake under iron deficiency can lead to a high silicate:nitrate ratio. The chemical composition of phytoplankton cells may also change with iron enrichment. Several environmental variables have been reported to affect diatom Si:C:N ratios¹⁴. However, very little is known of the effect of iron limitation on the diatom Si:N:P ratio.

To substantiate the field incubation results, nutrient consumption ratios were measured for Antarctic diatoms isolated from the Southern Ocean. Centric diatom *Chaetoceros dichchaeta* and pennate diatom *Nitzschia* sp. were grown under iron-deficient and iron-sufficient conditions in the laboratory, using filtered sea water collected from the Southern Ocean. In accordance with the results of field incubation experiments, the silicate:nitrate and silicate:phosphate consumption ratios of cultures grown under iron deficiency were twice as high as those of iron-sufficient cultures (Table 2). The laboratory culture experiments also demonstrated that iron-stressed, slowly growing diatoms produce high-nitrate, low-silicate

Table 2 Laboratory culture experiments of Antarctic diatom isolates

	Nutrient consumption ratio			Elemental composition			Cell volume (μm^3)
	Si(OH) ₄ :NO ₃ ⁻ (mol mol ⁻¹)	Si(OH) ₄ :PO ₄ ³⁻ (mol mol ⁻¹)	NO ₃ :PO ₄ ³⁻ (mol mol ⁻¹)	Si (pmol cell ⁻¹)	N (pmol cell ⁻¹)	P (pmol cell ⁻¹)	
<i>Chaetoceros dichaeta</i>							
Fe-deficient*	1.9	16	8.5	2.4	1.3	0.15	710
Fe-sufficient†	0.7	7.1	9.7	2.4	3.3	0.34	1200
<i>Nitzschia</i> sp.							
Fe-deficient*	2.1	42	20	3.6	1.8	0.09	3100
Fe-sufficient†	1.2	21	18	1.8	1.5	0.08	2700

* Initial concentration of total dissolved iron is 0.13 nmol l⁻¹.

† Initial concentration of total dissolved iron is 2.1 nmol l⁻¹.

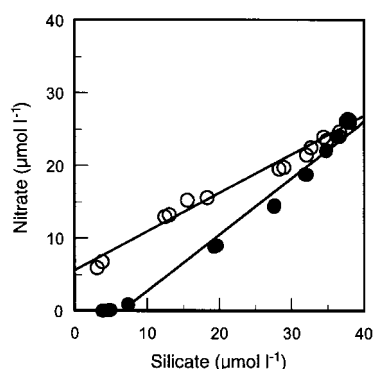


Figure 2 Nitrate versus silicate concentration in the seawater culture of Antarctic diatom *Chaetoceros dichaeta*. Open circles, iron-deficient conditions (0.13 nmol l⁻¹ Fe); filled circles, iron-sufficient conditions (2.1 nmol l⁻¹ Fe). The sea water for the culture was prepared from filtered sea water collected at the surface mixing layer of the Southern Ocean (65.4° S, 139.8° E). No major nutrient was added to the sea water, so the initial concentrations of nitrate and silicate reflect the *in situ* condition. Solid lines, slopes fitted by linear regression.

conditions, while iron-induced blooming of diatoms results in low-nitrate, high-silicate conditions (Fig. 2). The increases in silicate:nitrate and silicate:phosphate consumption ratios in *Nitzschia* sp. were due almost entirely to an increase of Si content per cell, whereas in *C. dichaeta* they were due to a decrease in N and P content per cell (Table 2). The reason for this different response between these two species is unknown, although I note that the structure of the silica deposition vesicle is different in cells from centric and pennate diatoms¹⁵. In *C. dichaeta*, a reduction in cell volume was also responsible for an increase in Si content per unit cell volume (Table 2). The reduction in cell volume causes an increase in the surface-to-volume ratio of the cell, and the amount of silica in a cell wall is proportional to the surface area of a cell¹⁶. Although very little is known of the role of iron in silicate uptake and silicate polymerization in diatom cells, these results suggest that the mechanisms regulating Si assimilation are different from those controlling assimilation of N and P.

Silicate dynamics is one of the important factors that regulate new production processes in the equatorial Pacific Ocean¹⁷ and other iron-limited waters¹⁸. The biological activity of carbon transport could be estimated from the distribution and flux of biogenic Si. Iron deficiency, however, appears to affect the diatom C:Si ratio in a similar manner to the Si:N and Si:P ratios, because iron-limited diatoms show little variability in C:N:P elemental composition⁷. Assuming the Redfield C:N:P ratio of 106:16:1, the measured Si:N ratios of *C. dichaeta* and *Nitzschia* sp. lead to C:Si ratios of 5.5–9.1 under iron-sufficient conditions and 3.3–3.6 under iron-deficient conditions. The estimated C:Si ratio of diatoms grown under iron-

sufficient conditions is nearly equal to the typical diatom C:Si ratio of 7.7 (ref. 14). These estimates suggest that twice as much carbon can be transported along with biogenic Si during iron-induced diatom blooming. The flux of iron may also affect the upper limit on the rate of new production in the equatorial upwelling zone, where silicate availability regulates the uptake and fate of both carbon and nitrate¹⁷.

Many hypotheses have been advanced to explain the glacial–interglacial change in atmospheric CO₂ concentration¹⁹. One is the iron hypothesis²⁰, which invokes an increased iron supply to the Southern Ocean from atmospheric dust during the glacial period. This would have stimulated phytoplankton production and biological pumping of carbon into the deep sea. There is palaeoceanographic evidence⁸ for greatly increased fluxes of biogenic detritus out of surface waters in the iron-fertilized, glacial Southern Ocean. However, Southern Ocean sedimentary records show a nonlinear relationship between organic carbon flux and opal accumulation rate⁸. This decoupling of organic carbon and opal records can be explained to some extent by the iron-induced change in diatom C:Si ratio observed in this study. As pointed out by Martin²¹, the change in C:Si ratio may also explain the failure²² to identify, in opal accumulation rates, evidence for greatly enhanced export production in the glacial Southern Ocean. The present findings in the modern ocean are consistent with palaeoceanographic records⁸ that support the iron hypothesis for the lower concentration of CO₂ in the glacial atmosphere. □

Methods

In vitro iron enrichment experiments. In the Southern Ocean (64.2° S, 140.7° E; cruise KH-94-4 in January 1995), the subarctic North Pacific (45° N, 165° E; cruise KH-93-4 in October 1993) and the equatorial Pacific²³ (0°, 159° W; cruise KH-93-4 in November 1993), unfiltered sea water, with its resident phytoplankton, was collected from a depth of 10–15 m using Teflon or silicone tubing (internal diameter, 12 mm) and a polypropylene pump system. This sea water was poured into 12-l polyethylene (cruise KH-94-4) or 1-l polycarbonate (cruise KH-93-4) acid-cleaned bottles. The bottles were either spiked with acidic iron chloride or left as controls, and were incubated in shipboard incubators at surface-water temperatures for 5–7 days. The incubators were covered with neutral-density screens to provide shading to 40% of the incident light level. In the light-limitation experiments in the Southern Ocean, blue-light filters were used with the screens to reduce the light intensity to 2.6% of the incident level. The treated water and the controls were sampled almost every other day for shipboard measurements of chlorophyll *a*, nutrients and total dissolvable iron. In case of the experiments with 1-l incubation bottles, replicate samples were taken from the replicate bottles and the bottles were not repetitively sampled. Sample treatments and subsampling procedures were carried out under class 100 clean-air conditions on board ship. Chlorophyll *a* samples were collected on Whatman GF/F filters, extracted into *N,N*-dimethylformamide and determined by fluorometry. Phytoplankton growth rates were calculated by linear regression from logarithmically transformed daily chlorophyll *a* data. Nutrients were measured by standard techniques using an auto-analyser. The difference in nutrient concentration between the first and

the last days of the experiments was used for the calculation of nutrient consumption ratios. Total dissolvable iron (the sum of dissolved and labile particulate iron at pH $\approx$ 3) was determined by preconcentration with chelating resin followed by a chemiluminescence detection method using an automatic iron analyser²⁴. Initial conditions are; 0.91 μg chlorophyll *a* l⁻¹, and 22.8, 1.24 and 18.7 $\mu\text{mol l}^{-1}$ of NO₃⁻, PO₄³⁻ and Si(OH)₄, respectively, in the Southern Ocean; 1.28 μg chlorophyll *a* l⁻¹, and 12.1, 1.15 and 22.0 $\mu\text{mol l}^{-1}$ of NO₃⁻, PO₄³⁻ and Si(OH)₄, respectively, in the subarctic North Pacific; 0.48 μg chlorophyll *a* l⁻¹, and 3.7, 0.42 and 2.1 $\mu\text{mol l}^{-1}$ of NO₃⁻, PO₄³⁻ and Si(OH)₄, respectively, in the equatorial Pacific.

Laboratory culture experiments. Strains of *Chaetoceros dichæta* and *Nitzschia* sp. were isolated from the marginal ice zone (63.8–65.5° S) of the Australian sector in the Southern Ocean. Cells were pre-acclimatized at low Fe concentration (0.13 nmol l⁻¹) for one month, and grown for three weeks at 2 $\pm$ 0.5 °C for *C. dichæta* and 4 $\pm$ 0.5 °C for *Nitzschia* sp. under continuous light (22 $\mu\text{E m}^{-2} \text{s}^{-1}$) in filtered (0.03- μm pore size polyethylene hollow-fibre filter) Southern Ocean sea water using sterile, 1-l acid-washed polycarbonate bottles. This sea water had been collected from 15 m depth at the marginal ice zone (65.4° S, 139.8° E) using a trace-metal-clean pump system, and had been stored in the dark at 1 °C for 1.5 yr. The sea water contained 27 $\mu\text{mol l}^{-1}$ NO₃⁻, 1.7 $\mu\text{mol l}^{-1}$ PO₄³⁻, 46 $\mu\text{mol l}^{-1}$ Si(OH)₄ and 0.13 nmol l⁻¹ total dissolvable Fe. Trace metals except for iron (10 nmol l⁻¹ MnCl₂, 10 nmol l⁻¹ ZnCl₂, 10 nmol l⁻¹ CoCl₂, 1 nmol l⁻¹ CuCl₂, 1 nmol l⁻¹ (NH₄)₆Mo₇O₂₄), 90 nmol l⁻¹ Na₂EDTA and vitamins (30 nmol l⁻¹ thiamine hydrochloride, 0.2 nmol l⁻¹ biotin, and 0.04 nmol l⁻¹ vitamin B₁₂) were added to the sea water. Iron-sufficient cultures were also enriched with 2 nM FeEDTA, while iron-deficient cultures received no additional iron. Sterile techniques were used, and bacteria were never apparent in the cultures. Subsamples were collected every day for measurements of nutrients, chlorophyll *a* and cell number. All procedures were carried out under class 100 clean-room conditions at 1 °C. Nutrients were measured by standard techniques using an auto-analyser. The nutrient consumption ratio was calculated from the change in nutrient data in exponential-phase batch cultures. Estimates of particulate N, P and Si were made by a mass-balance method, which assumes that a negligible amount of the inorganic nutrient in the culture sea water is converted into dissolved organic compounds¹¹. The decrease in each nutrient concentration during the exponential growth phase was divided by the increase in cell number (which had been determined using an inverted microscope) to yield cellular elemental composition of N, P and Si. Cell volumes were calculated from microscopic measurements of cell diameter and length for 50 randomly selected cells. The two species of diatoms showed significantly lower growth rates in iron-deficient cultures than did those in iron-sufficient cultures.

Received 21 November 1997; accepted 6 May 1998.

- Martin, J. H. & Fitzwater, S. E. Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic. *Nature* **331**, 341–343 (1988).
- Martin, J. H. *et al.* Testing the iron hypothesis in ecosystems of the equatorial Pacific. *Nature* **371**, 123–129 (1994).
- Coale, K. H., Fitzwater, S. E., Gordon, R. M., Johnson, K. S. & Barber, R. T. Control of community growth and export production by upwelled iron in the equatorial Pacific Ocean. *Nature* **379**, 621–624 (1996).
- Martin, J. H., Fitzwater, S. E. & Gordon, R. M. Iron deficiency limits phytoplankton growth in Antarctic waters. *Glob. Biogeochem. Cycles* **4**, 5–12 (1990).
- de Baar, H. J. W. *et al.* Importance of iron for phytoplankton blooms and carbon dioxide drawdown in the Southern Ocean. *Nature* **373**, 412–415 (1995).
- Price, N. M., Ahner, B. A. & Morel, F. M. M. The equatorial Pacific Ocean: Grazer-controlled phytoplankton populations in an iron-limited ecosystem. *Limnol. Oceanogr.* **39**, 520–539 (1994).
- Green, R. M., Geider, R. J. & Falkowski, P. G. Effect of iron limitation on photosynthesis in a marine diatom. *Limnol. Oceanogr.* **36**, 1772–1782 (1991).
- Kumar, N. *et al.* Increased biological productivity and export production in the glacial Southern Ocean. *Nature* **378**, 675–680 (1995).
- Duce, R. A. & Tindale, N. W. Atmospheric transport of iron and its deposition in the ocean. *Limnol. Oceanogr.* **36**, 1715–1726 (1991).
- Paasche, E. Silicon content of five marine plankton diatom species measured with a rapid filter method. *Limnol. Oceanogr.* **25**, 474–480 (1980).
- Harrison, P. J., Conway, H. L., Holmes, R. W. & Davis, C. O. Marine diatoms grown in chemostats under silicate or ammonium limitation. III. Cellular chemical composition and morphology of *Chaetoceros debilis*, *Skeletonema costatum*, and *Thalassiosira gravida*. *Mar. Biol.* **43**, 19–31 (1977).
- Hutchins, D. A. & Bruland, K. W. Iron-limited diatom growth and Si:N uptake ratios in a coastal upwelling. *Nature* **393**, 561–564 (1998).
- Raven, J. A. The iron and molybdenum use efficiencies of plant growth with different energy, carbon and nitrogen sources. *New Phytol.* **109**, 279–287 (1988).
- Brzezinski, M. A. The Si:C:N ratio of marine diatoms: Interspecific variability and the effect of some environmental variables. *J. Phycol.* **21**, 347–357 (1985).
- Round, F. E., Crawford, R. M. & Mann, D. G. *The Diatoms* (Cambridge Univ. Press, 1990).

- Paasche, E. The influence of cell size on growth rate, silica content, and some other properties of four marine diatom species. *Norw. J. Bot.* **20**, 197–204 (1973).
- Dugdale, R. C. & Wilkerson, F. P. Silicate regulation of new production in the equatorial Pacific upwelling. *Nature* **391**, 270–273 (1998).
- Dugdale, R. C., Wilkerson, F. P. & Minas, H. J. The role of a silicate pump in driving new production. *Deep-Sea Res.* **1**, 42, 697–719 (1995).
- Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, C. Vostok ice core provides 160,000-year record of atmospheric CO₂. *Nature* **329**, 408–414 (1987).
- Martin, J. H. Glacial-interglacial CO₂ change: The iron hypothesis. *Paleoceanography* **5**, 1–13 (1990).
- Martin, J. H. in *Primary Productivity and Biogeochemical Cycles in the Sea* (eds Falkowski, P. G. & Woodhead, A. D.) 123–137 (Plenum, New York, 1992).
- Mortlock, R. A. *et al.* Evidence for lower productivity in the Antarctic Ocean during the last glaciation. *Nature* **351**, 220–222 (1991).
- Takeda, S. & Obata, H. Response of equatorial Pacific phytoplankton to subnanomolar Fe enrichment. *Mar. Chem.* **50**, 219–227 (1995).
- Obata, H., Karatani, H., Matsui, M. & Nakayama, E. Fundamental studies for chemical speciation of iron in seawater with an improved analytical method. *Mar. Chem.* **56**, 97–106 (1997).

Acknowledgements. I thank H. Ogawa, T. Usui, and S. Watanabe for shipboard assistance in processing nutrient samples; H. Obata for assistance in processing iron samples; K. Watanabe for Antarctic phytoplankton samples; and I. Koike, K. Kawaguchi, and M. Terasaki for their support during the cruises. My appreciation is also extended to the crew and officers of the RV *Hakuho-maru*.

Correspondence and requests for materials should be addressed to S.T. (e-mail: s-takeda@criepi.denken.or.jp).

Oxygen isotope evidence for slab-derived fluids in the sub-arc mantle

John M. Eiler*, Brent McInnes†, John W. Valley‡, Colin M. Graham§ & Edward M. Stolper*

* Division of Geological and Planetary Science, California Institute of Technology, Mail Code 170-25, Pasadena, California 91125, USA

† CSIRO Exploration and Mining, PO Box 136, North Ryde, NSW 2113, Australia

‡ Department of Geology and Geophysics, University of Wisconsin, Madison, Wisconsin 53706, USA

§ Department of Geology and Geophysics, University of Edinburgh, Edinburgh, UK

The subduction of oceanic lithosphere is thought to enrich the mantle in elements concentrated in altered oceanic crust and its sedimentary cover (for example, H₂O, CO₂ and alkalis)^{1,2}. This enrichment is generally inferred from the geochemistry of island-arc lavas³. More direct evidence—such as samples from the mantle with a clear crustal origin⁴—is rare. Inclusions of silicate glass within mantle-derived minerals can have major- and trace-element compositions unlike basalt^{5,6}, and sometimes contain ‘enriched’ isotopic compositions of Sr, Nd and Pb, suggesting that the inclusions are partial melts of subducted oceanic crust or sediments⁶. Alternatively, some of these alkali-rich inclusions may have been produced by melting peridotites to low degrees (possibly in the presence of volatiles)⁷. Here we present oxygen isotope data from silicate glass inclusions obtained from mantle olivine samples in an island-arc setting. These data provide direct evidence that the inclusions are derived from a source rich in material from the subducted oceanic crust and therefore that slab-derived fluids have infiltrated the sub-arc mantle wedge.

Most mantle-derived magmas and ultramafic rocks have well-defined $\delta^{18}\text{O}$ values of $5.7 \pm 0.5\%$, where $\delta^{18}\text{O} = ((^{18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{18}\text{O}/^{16}\text{O})_{\text{SMOW}} - 1) \times 1,000$ and $(^{18}\text{O}/^{16}\text{O})_{\text{SMOW}} = 0.0020052$ (SMOW being standard mean ocean water) (refs 8–10). In contrast, altered basaltic and gabbroic rocks of the oceanic crust range from 0‰ to +12‰ (ref. 11), clastic sediments are typically +12‰ to +20‰ (ref. 12), and carbonate and silicate pelagic sediments are typically +20‰ to +25‰ (ref. 13). Consequently, the isotopic composition of oxygen in a mantle-derived mineral or melt is a potentially sensitive indicator of subducted crustal materials as its protolith or in its source.

Alkali basalts from Simberi island, New Guinea, contain lherzo-

lite and harzburgite xenoliths and xenocrysts derived from the mantle above the Manus–Kilnailau subduction zone^{14,15}. These xenoliths and xenocrysts are composed predominantly of olivine, orthopyroxene, clinopyroxene and spinel. They contain secondary minerals (including Na-rich clinopyroxene, amphibole, Ti-rich phlogopite, magnetite, K-rich feldspar, calcite, apatite and anhydrite), rare secondary-vapour inclusions, and even rarer patches of an Na-rich silicate glass, either as inclusions in minerals or as domains in veins of secondary minerals (Fig. 1; both textural types of glass will be referred to as ‘inclusions’). Textural and chemical evidence suggest that the secondary minerals were produced by reaction of Na-rich silicate melts (or perhaps fluids) with the host minerals, accompanied by precipitation of daughter minerals from the melt¹⁴. A particularly large vein of secondary Na-rich pyroxene and glass (3.5 mm × 100–600 µm) was found in a single olivine crystal (1.5 mm × 4 mm; sample 310 from ref. 14). The olivine has a chemical composition indicating that it is a mantle xenocryst (Fo₉₁, 0.11 weight per cent (wt%) CaO, 2,000 p.p.m. Ni, compared with Fo_{77–83}, 0.23 wt% CaO, <1,300 p.p.m. Ni in phe-

nocrysts crystallized from the host lava). This vein includes two regions of glass (Fig. 1) suitable for multiple analyses by secondary ion mass spectrometry (SIMS), which we used to analyse the oxygen isotope composition and water content of the glass in these two regions; we also measured the concentrations of major and trace elements in the glasses (and associated secondary minerals) in these regions and in a related inclusion of intimately mixed glass and carbonate daughter crystals in another olivine xenocryst from the same sample. Analytical methods are summarized at the end of this Letter and results are presented in Table 1 and Fig. 2.

The silicate glass inclusions are phonolitic (that is, 78% normative feldspar, 18% nepheline): similar to some silicate glass inclusions in mantle xenoliths from oceanic hot spots and the continental lithosphere^{5,6}, but more alkaline than previously described inclusions in arcs⁵. Four analyses of $\delta^{18}\text{O}$ in inclusion A yielded an average value of $11.3 \pm 1.3\text{‰}$ (1σ). This inclusion is homogeneous at the precision of our analyses. Two analyses of inclusion B yielded values of 8.8‰ and 9.2‰; these are also indistinguishable given our precision. The olivine grain that hosts these inclusions was analysed five times, averaging $5.7 \pm 0.8\text{‰}$. This value is within analytical uncertainty of the range typical of mantle-derived olivine determined by fluorination analyses ($5.2 \pm 0.2\text{‰}$) (ref. 10). We used conventional fluorination (with an uncertainty of $\pm 0.1\text{‰}$) (ref. 16) to measure the $\delta^{18}\text{O}$ of Na-rich clinopyroxene separates (sub-samples 310-4, 310-5, and 310-7) from a metasomatized lherzolite xenolith from the same sample. Three analyses yielded values of 6.2‰, 6.9‰ and 10.3‰ — all enriched in ^{18}O relative to typical mantle pyroxenes ($5.6 \pm 0.2\text{‰}$) (ref. 10).

The trace element contents of the phonolitic glasses have several distinctive features (Fig. 2a): (1) enrichments in Rb and B relative to mid-ocean-ridge basalts (MORB), (2) depletions in Ba, Nb, Zr, Ti and Li relative to MORB, and (3) light-rare-earth abundances comparable to MORB, coupled with depletion in heavy rare earths. The depletion of Zr in these inclusions is particularly pronounced. The trace-element contents of associated metasomatic minerals and of the carbonate-rich inclusion are broadly similar to those of the phonolitic glasses, although they differ in detail (for example, the carbonate-rich inclusion is higher in Sr and Ba; Fig. 2a). The phonolitic glasses are water-rich: $7.9 \pm 0.1\text{ wt\% H}_2\text{O}$.

The Sr, Ba, Li, Rb, K, and to a smaller extent Ti concentrations in the phonolitic glasses have probably been modified by crystallization of (and/or interaction with) phases such as calcite, apatite, titanite, phlogopite, Na-rich clinopyroxene, and K-feldspar, all of which are present in the vein assemblages. Compositions of the last three phases are given in Table 1 and plotted in Fig. 2a for reference. We have considered the trace-element compositions of the glasses and these related minerals in the discussion below of possible origins of the infiltrating melts. The water content of the glasses have probably been only modestly increased by crystallization of hydrous daughter minerals.

The $\delta^{18}\text{O}$ values determined for inclusions A and B are 3–5‰ higher than those for common mafic mantle melts⁸ and 2–4‰ higher than those expected for more siliceous melts in oxygen isotope equilibrium with mantle peridotite^{10,17}. The larger of the two inclusions (A) is more extreme in this respect than the smaller (B). Several lines of evidence suggest that the ^{18}O -enrichments we observe in the glass inclusions reflect the pre-entrapment $\delta^{18}\text{O}$ values of the silicate liquids from which they were formed rather than post-entrapment enrichment.

(1) The inclusions analysed do not appear to be devitrified.

(2) Cl concentrations are $\leq 0.1\text{ wt\%}$ in the inclusions analysed, compared with $\sim 3\%$ in palagonite formed by alteration of glass in the groundmass of the lava that hosts the xenocrysts we have studied¹⁴. The glass inclusions therefore appear not to have undergone the alteration experienced by groundmass glass in the same hand specimen.

(3) Values of $\delta^{18}\text{O}$ and the abundances of B and H_2O are

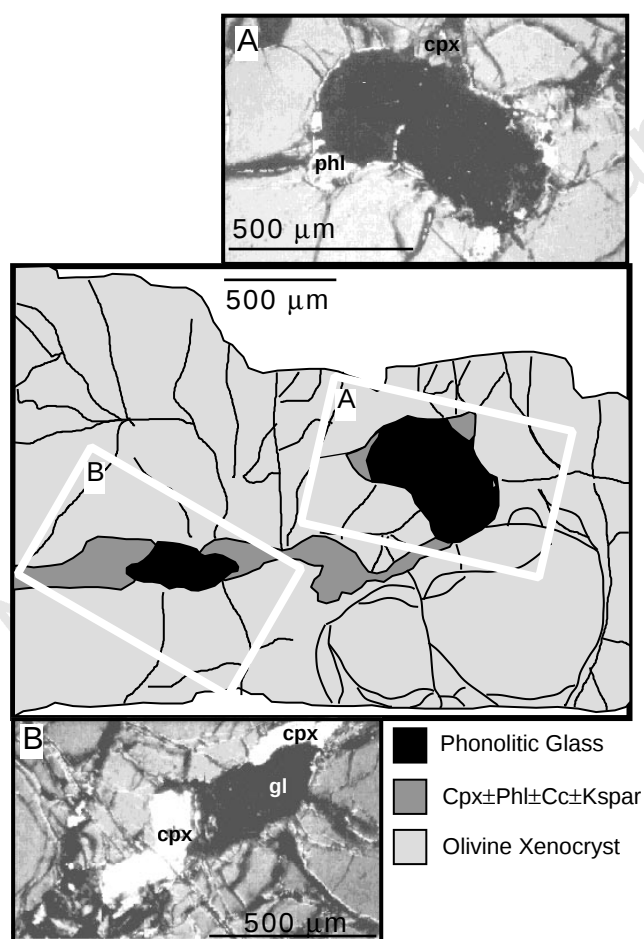


Figure 1 Transmitted light photomicrographs (A and B), taken with crossed polarizers, of regions containing the phonolitic glasses analysed in this study. Glass inclusions are hosted by an olivine xenocryst in sample 310 from ref. 14. The glass in region A is relatively free of secondary and daughter minerals, whereas that in region B is associated with secondary Na-rich clinopyroxene and thus possibly records significant reaction between silicate melt and host olivine. Black regions are phonolitic glass (gl); bright areas are Na-rich clinopyroxene (cpx) and phlogopite (phl); and the intermediate grey is host olivine. The central figure is a tracing of the outline of the host olivine xenocryst, inclusion phases, and cracks. Cc, calcite; Kspar, K-feldspar.

reproducible to within analytical precision for each inclusion (two to four analyses each), whereas they are usually heterogeneous in altered minerals and glasses^{18,19}.

(4) High values of $\delta^{18}\text{O}$ in secondary Na-rich clinopyroxene (which is more resistant to low-temperature alteration than glass) in a xenolith from the same lava provide corroborating evidence that the infiltrating melt that metasomatized these xenoliths was ^{18}O -enriched, in agreement with our observations.

For these reasons we interpret the measured $\delta^{18}\text{O}$ values as primary (that is, inherited from the silicate melts from which the glasses formed). The best estimate of the $\delta^{18}\text{O}$ of these liquids before

interaction with the host olivine is the average for glass in inclusion A (11.3‰), because inclusion B is associated with more secondary Na-rich clinopyroxene and is thus more likely to have decreased in $\delta^{18}\text{O}$ by reaction with host olivine. Such a large enrichment in ^{18}O relative to MORB and mantle peridotites suggests that a large fraction (20–100%) of the oxygen in the sources of these phonolitic melts was derived from hydrothermally altered basaltic rocks, sediments, or fluids equilibrated with such materials. Given that these glasses occur in a mantle-derived mineral from a volcano overlying a recently deactivated subduction zone^{14,15}, subduction is the most likely mechanism for bringing such ^{18}O -enriched crustal

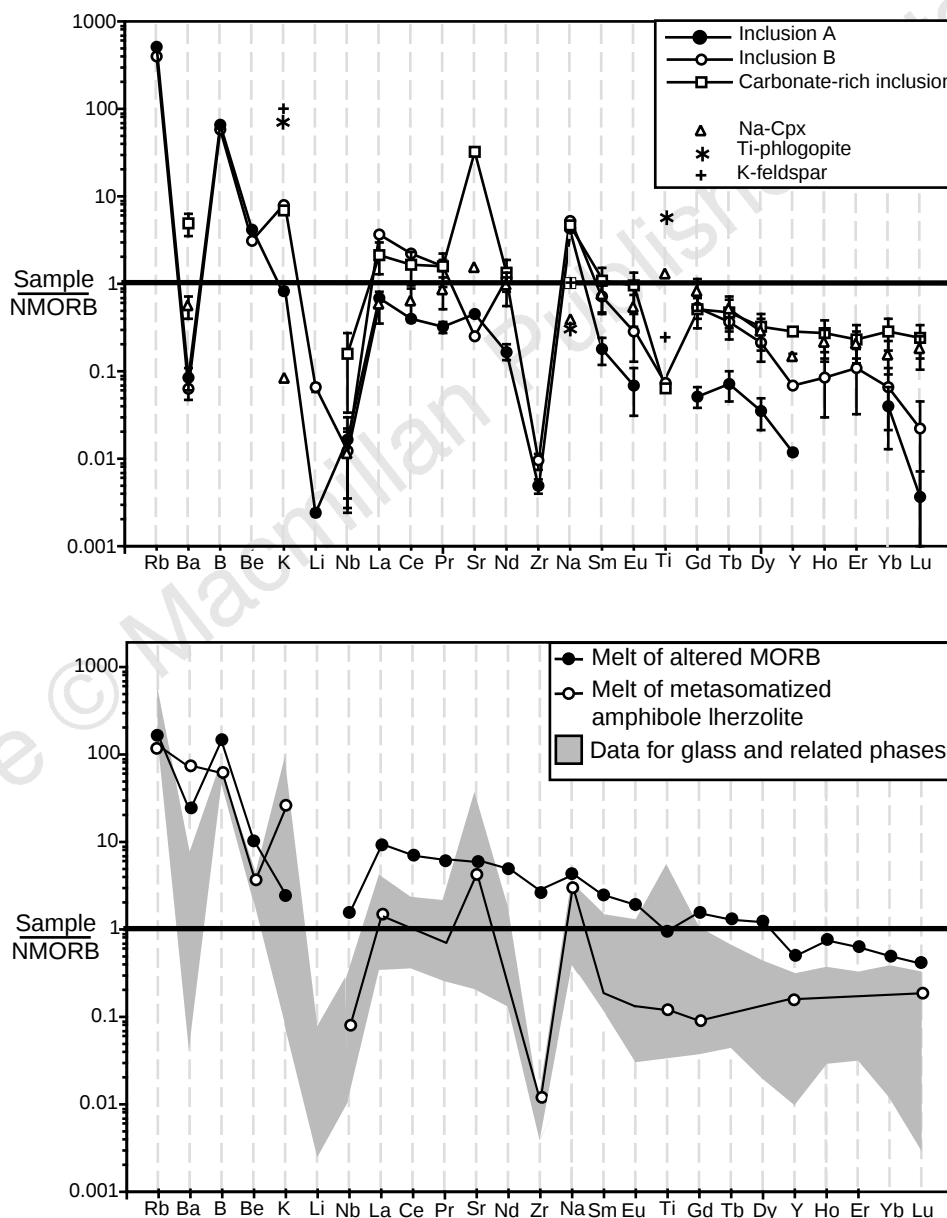


Figure 2 Abundances of trace and minor elements in glasses and minerals analysed in this study and comparison with models for the origin of phonolitic melt. **a**, Concentrations of trace elements in phonolitic glass and related phases, normalized to concentrations in NMORB^{2,35} (here NMORB indicates normal MORB). Elements are listed from left to right in approximate increasing order of solid/melt partition coefficients^{2,21}. **b**, Comparison of calculated trace element concentrations for low-degree partial melts of subducted basalt (filled symbols) and of metasomatized amphibole lherzolite (unfilled symbols) with observed abundances in phonolitic glass and related phases (grey field). This field

encompasses the data for daughter and secondary minerals shown in **a** because of the possibility that they represent significant components of the infiltrating melt. No model calculations were made for Li owing to the relatively poor constraints on partitioning and concentrations in altered MORB. Low Ba contents in phonolitic glasses are not well explained by any models using average subducted crustal compositions and suggest either a residual Ba-rich phase in the source of phonolitic melt or an unusually Ba-poor subducted component (see ref. 36).

Table 1 Chemical and isotopic data for phonolitic glass and related phases

Component	Inclusion A	Inclusion B	Olivine host (incl. A + B)	Secondary Ti-phlogopite (incl. A)	Secondary K-feldspar (incl. A)	Secondary Na-clinopyroxene (incl. B)	Mixed carbonate + glass inclusion*
(wt%)							
SiO ₂	56.4	53.4	41.1	37.85	62.00	48.9	
TiO ₂	b.d.l.	0.09	b.d.l.	7.41	0.30	1.76	
Al ₂ O ₃	24.30	24.29	b.d.l.	13.12	19.80	5.60	
FeO†	0.10	0.37	8.77	5.37	0.30	4.13	
MgO	0.02	b.d.l.	50.67	20.4	0.10	14.47	
NiO			0.25				
CaO	1.35	2.21	0.11	0.08	1.30	22.48	
Na ₂ O	9.30	10.21	b.d.l.	0.85	2.80	1.37	
K ₂ O	0.10	1.01	b.d.l.	9.47	12.60	0.01	
H ₂ O	7.8	7.9	b.d.l.			b.d.l.	
Cl	0.02	0.02			0.03		
(p.p.m.)‡							
Rb	290	219				3.50	31.27
Ba	0.525	0.41					
B	40.3	35.0					
Be	1.5	1.1					
Li	0.012	0.33					
Nb	0.038	0.03				0.03	0.36
La	1.70	9.03				1.46	5.25
Ce	2.94	16.63				4.76	12.22
Pr	0.42	2.13				1.12	2.05
Sr	39.9	22.4				140	2922
Nd	1.23	8.21				6.91	9.81
Zr	0.362	0.70					
Sm	0.47	1.88				1.97	2.90
Eu	0.07	0.30				0.55	0.99
Gd	0.19	2.00				3.04	1.90
Tb	0.05	0.25				0.35	0.32
Dy	0.16	0.98				1.31	1.46
Y	0.33	1.93				4.10	8.13
Ho	b.d.l.	0.09				0.22	0.28
Er	b.d.l.	0.32				0.60	0.70
Yb	0.12	0.20				0.47	0.88
Lu	0.002	0.01				0.08	0.11
δ ¹⁸ O _{SMOW} (±1‰)	12.2, 9.1, 11.7, 12.1	8.8, 9.2	5.6, 7.0, 4.5, 5.6, 5.8			6.2, 6.9, 10.3§	
Average	11.3	9.0	5.7				
External error	1.1	0.9	0.5				

Blank spaces indicate elements or oxygen isotope ratios that were not determined in that phase. b.d.l., below detection limit.

* Measurements of ~30-μm spots integrating carbonate and glass composition.

† Total Fe calculated as FeO.

‡ 2σ errors: 1–6% for >10 p.p.m. elements, 3–20% for 1–10 p.p.m. elements, 10–90% for <1 p.p.m. elements. Trace element data are averages of two to four measurements. Major element data are averages of multiple measurements.

§ All measurements of pyroxene are fluorination of mineral separates, ±0.1‰ precision; see ref. 16 for methods.

|| Cumulative uncertainty, including the uncertainty in the mean of sample measurements, reproducibility of the working standard, and reproducibility of the differences in instrumental fractionation between standards²⁴.

rocks or fluids into the mantle. To our knowledge, this study is the first direct demonstration of the existence of mantle fluids or melts with strongly elevated δ¹⁸O values.

The oxygen isotope evidence for an affinity to subducted crustal material cannot distinguish between an origin of the phonolitic melts by: (1) direct melting of hydrothermally altered basalt and/or sediment, or (2) by melting of peridotite metasomatized by fluid or melt derived from the subducted slab. The first of these hypotheses is the simplest and perhaps most intuitively appealing, but it is not easily reconciled with the trace element concentrations of the glasses and associated minerals (Fig. 2). In particular, the concentrations of Nb, the rare earth elements and especially Zr are significantly lower than expected in a melt of a crustal rock or sediment. This is illustrated by a comparison of the data with the predicted trace-element concentration in a low-degree melt of altered MORB (Fig. 2b; calculated as a 5% batch modal melt of an amphibole-bearing eclogite [clinopyroxene₄₀garnet_{24.5}amphibole_{24.5}rutile₂] using concentrations and partitioning data from refs 20–22 and from references in ref. 21). The low observed concentrations of these elements are better explained by a depleted, peridotitic source (see, for example, refs 23, 24). Results of one model that could fit most of the observed trace element compositions are shown in Fig. 2b: phonolites are described as a 5% batch modal melt of a highly depleted peridotite to which the trace elements from a hydrous fluid in equilibrium with MORB have been added (assuming a lherzolite

source [olivine₆₇orthopyroxene₁₅amphibole₁₃clinopyroxene₃spinel₂]; taking concentrations and partition coefficients from refs 2, 21, 23–29 and from references in ref. 21; and assuming 0.1 wt% Na₂O and no Rb, Ba, K, B or Be in the lherzolite before metasomatism). Models differing from this in detail are also permissible, but all successful models examined by us are similar to this example; otherwise, we have not been able to fit the restriction of exceptionally low Zr concentration and the low concentrations of incompatible elements generally. The trace-element data therefore favour an origin of the ¹⁸O-enriched phonolitic melts as low-degree partial melts of a highly depleted peridotite that was metasomatized by slab-derived fluids. This hypothesis is consistent with the high observed B/Be ratio in these glasses, diagnostic of slab-derived fluids but not expected of slab-derived melts^{1,30}, and with experimental studies indicating that melts with similar compositions can be derived by melting that leaves a harzburgitic residue³¹.

There are no known mantle peridotites with values of δ¹⁸O as high as those that we infer for the sources of the phonolitic glasses we have studied, nor has the existence of such peridotites been inferred previously from the oxygen isotope geochemistry of basaltic arc-related lavas. However, sub-lithospheric xenoliths are rare in arc settings, so even if such ¹⁸O-enriched peridotites were common in mantle wedges, we would not necessarily expect them to be represented in xenolith collections. The fact that island arc lavas are not as enriched in ¹⁸O as the phonolitic glasses from this study

suggests that the sources of high- $\delta^{18}\text{O}$, phonolitic melts are but a small component of the sub-arc mantle or that, if common, they are not sufficiently fertile to contribute significantly to typical island-arc basalts (for example, if they are always as depleted as inferred in this study). Similarly, if the most highly metasomatized peridotites in the mantle wedge are, as expected, just above the subducting slab, they might be too cool to participate extensively in basalt generation, although they might in special circumstances become hot enough to produce phonolites and other siliceous melts (which require lower temperatures than needed to produce basaltic lavas⁷). We conclude that there is no observational basis to discount the existence of highly depleted, extensively metasomatized peridotites in the sub-arc mantle and the subsequent melting of such sources to produce high $\delta^{18}\text{O}$ phonolites. Such a two-stage process is widely accepted as the mechanism by which the geochemical characteristics of slab-derived fluids are transported into the sub-arc mantle and then into igneous rocks produced in these environments^{2,30,32,33}. If high- $\delta^{18}\text{O}$ phonolitic liquids and/or their inferred highly depleted, high- $\delta^{18}\text{O}$ peridotitic sources contribute to the sources of common arc-related basalts (as opposed to being the product of processes restricted to the anomalous tectonic setting of the Simberi island volcanics^{14,15}), they might be detected in the geochemical variability of such lavas. For example, the phonolites that we have studied are distinctive in their depletion in certain incompatible elements (Zr, Nb, rare-earth elements), and thus our study suggests that future work should investigate island-arc basalts for an association of depletion in these elements with enrichment in ^{18}O . □

Methods

$^{18}\text{O}^-/^{16}\text{O}^-$ ratios were measured by secondary-ion mass spectrometry (SIMS) with the Cameca ims-4f at the University of Edinburgh, by using an 8 nA primary beam of $^{133}\text{Cs}^+$ with an impact energy of 14.15 keV and an energy offset for secondary ions of 350 ± 25 eV. Each analysis comprised 100 cycles of alternating measurements of $^{18}\text{O}^-$ (5 s counting) and $^{16}\text{O}^-$ (1 s counting). Typical count rates were 10^6 c.p.s. on $^{16}\text{O}^-$ and 2×10^3 c.p.s. on $^{18}\text{O}^-$ for an expected precision in $^{18}\text{O}^-/^{16}\text{O}^-$ of $\pm 1\%$, 1σ . Point-to-point precision on multiple analyses of standards was $\pm 0.9\%$. Analyses of phonolitic glass and olivine were standardized against Amelia albite analysed in the same analytical session and previously established differences in instrumental fractionation between Amelia albite, olivines and various Na–Al-silicate glasses (including a hydrous glass)³⁴. Measurement of a secondary standard of jadeitic glass in the same analytical session and standardized against the same body of data for Amelia albite resulted in a $\delta^{18}\text{O}$ value of $13.3 \pm 0.6\%$ —within precision of the accepted value of 12.7% based on fluorination analyses. Although the $\delta^{18}\text{O}$ of the olivine grain analysed in this study is not known independently, the agreement of our measurement of this sample with the known range of mantle olivines is also consistent with this demonstration of the accuracy of our ion-probe analyses.

$\text{H}^-/^{28}\text{Si}^-$ ratios were measured with the same instrumentation and conditions as for oxygen isotope ratio measurements. These analyses were standardized with nominally anhydrous albitic glass and crystalline albite to establish the background, and albitic glasses containing 1 wt%, 3 wt% and 7 wt% H_2O to establish a calibration between the measured $\text{H}^-/^{28}\text{Si}^-$ ratio and water content. This calibration was repeated on three different days and found to have an uncertainty of $\pm 10\%$, relative, both in the regression of a working curve for any given day and in the day-to-day reproducibility of the working curve. Point-to-point reproducibility on each standard on any given day was significantly better than this (± 0.5 – 1.0% , relative). The background was negligible relative to the water contents of our unknowns.

Concentrations of trace elements in silicate glass inclusions were also made by SIMS at the University of Edinburgh, with an 8 nA primary beam of O^- with an impact energy of 14.4 keV and an energy offset for secondary ions of 75 ± 25 eV. Measurements were based on a comparison of the ratio of the intensity of each element to that of $^{28}\text{Si}^+$, were standardized relative to NBS-610 glass and previously established working curves, and were checked by the analysis of BIR-1 glass as an unknown. Concentrations of trace elements in the mixed carbonate plus glass inclusion and Na-rich clinopyroxene were deter-

mined at the California Institute of Technology with similar methods but normalizing relative to $^{40}\text{Ca}^+$ rather than to $^{28}\text{Si}^+$. The measurements for the mixed carbonate plus glass inclusion had a large enough spot size for the analysis to integrate the compositions of both intimately mixed phases. All reported data in Table 1 and Fig. 2a are averages of two to four measurements. Error bars in Fig. 2a indicate $\pm 2\sigma$ uncertainties based on counting statistics. Concentrations of major elements in silicate glasses and minerals were determined by electron microprobe at California Institute of Technology. Analytical details are described in ref. 14.

Received 24 September 1997; accepted 16 March 1998.

1. Leeman, W. P., Carr, M. J. & Morris, J. D. Boron geochemistry of the Central American volcanic arc: constraints on the genesis of subduction-related magmas. *Geochim. Cosmochim. Acta* **58**, 149–168 (1994).
2. Stolper, E. M. & Newman, S. The role of water in the petrogenesis of Mariana trough magmas. *Earth Planet. Sci. Lett.* **121**, 293–325 (1994).
3. Morris, J. & Tera, F. ^{10}Be and ^9Be in mineral separates and whole rocks from volcanic arcs: implications for sediment subduction. *Geochim. Cosmochim. Acta* **53**, 3197–3206 (1989).
4. Daniels, L. R. M., Gurney, J. J. & Harte, B. A crustal mineral in a mantle diamond. *Nature* **379**, 153–156 (1996).
5. Schiano, P. & Clocchiatti, R. Worldwide occurrence of silica-rich melts in sub-continental and sub-oceanic mantle minerals. *Nature* **368**, 621–624 (1994).
6. Hauri, E. H., Schimizu, N., Dieu, J. J. & Hart, S. R. Evidence for hot-spot-related carbonatite metasomatism in the oceanic upper mantle. *Nature* **365**, 221–227 (1993).
7. Baker, M. B., Hirschmann, M. M., Ghiorso, M. S. & Stolper, E. M. Compositions of near-solidus peridotite melts from experiments and thermodynamic calculations. *Nature* **375**, 308–311 (1995).
8. Ito, E., White, W. M. & Gopel, C. The O, Sr, Nd and Pb isotope geochemistry of MORB. *Chem. Geol.* **62**, 157–176 (1987).
9. Eiler, J. M. *et al.* Oxygen isotope variations in ocean island basalt phenocrysts. *Geochim. Cosmochim. Acta* **61**, 2281–2293 (1997).
10. Mathey, D., Lowry, D. & Macpherson, C. Oxygen isotope composition of mantle peridotite. *Earth Planet. Sci. Lett.* **128**, 231–241 (1994).
11. Muehlenbachs, K. in *Stable Isotopes in High Temperature Geological Processes* (eds Valley, J. W., Taylor, H. P. & O'Neill, J. R.) 425–444 (Mineralogical Society of America, Washington DC, 1986).
12. Arthur, M. A., Anderson, T. F. & Kaplan, I. R. in *SEPM Short Course* no. 10 (SEPM, Tulsa, 1986).
13. Kolodny, Y. & Epstein, S. Stable isotope geochemistry of deep sea cherts. *Geochim. Cosmochim. Acta* **40**, 1195–1209 (1976).
14. McInnes, B. I. A. & Cameron, E. M. Carbonated, alkaline hybridizing melts from the sub-arc environment: mantle wedge samples from the Tabar–Lihir–Tanga–Feni arc, Papua New Guinea. *Earth Planet. Sci. Lett.* **122**, 125–141 (1994).
15. Taylor, B. Bismarck sea: evolution of a back-arc basin. *Geology* **7**, 171–174 (1979).
16. Taylor, H. P. Jr. & Epstein, S. Relationship between $^{18}\text{O}/^{16}\text{O}$ ratios in coexisting minerals of igneous and metamorphic rocks. *Geol. Soc. Am. Bull.* **73**, 461–480 (1962).
17. Palin, J. M., Epstein, S. & Stolper, E. M. Oxygen isotope partitioning between rhyolitic glass/melt and CO_2 : an experimental study at 550–950 °C and 1 bar. *Geochim. Cosmochim. Acta* **60**, 1963–1973 (1996).
18. Eiler, J. M., Valley, J. W., Graham, C. M. & Baumgartner, L. P. Ion microprobe evidence for the mechanisms of stable isotope retrogression in high-grade metamorphic rocks. *Contrib. Mineral. Petrol.* **118**, 365–378 (1995).
19. Spivack, A. J. & Edmond, J. M. Boron isotope exchange between seawater and the oceanic crust. *Geochim. Cosmochim. Acta* **51**, 1033–1043 (1987).
20. McCulloch, M. T. & Gamble, J. A. Geochemical and geodynamical constraints on subduction zone magmatism. *Earth Planet. Sci. Lett.* **102**, 358–374 (1991).
21. Green, T. H. Experimental studies of trace-element partitioning applicable to igneous petrogenesis—Sedonia 16 years later. *Chem. Geol.* **117**, 1–36 (1994).
22. Staudigel, H., Plank, T., White, W. & Schminke, H. U. Geochemical fluxes during seafloor alteration of the basaltic upper oceanic crust: DSDP sites 417 and 418 (overview). *AGU Geophys. Monogr.* **96**, 19–38 (1996).
23. Salters, V. J. M. & Shimizu, N. World-wide occurrence of HFSE-depleted mantle. *Geochim. Cosmochim. Acta* **52**, 2177–2182 (1988).
24. Sharma, M. & Wasserburg, G. J. The neodymium isotopic compositions and rare earth patterns in highly depleted ultramafic rocks. *Geochim. Cosmochim. Acta* **60**, 4537–4550 (1996).
25. Johnson, K. T. M., Dick, H. J. B. & Shimizu, N. Melting in the oceanic upper mantle: an ion microprobe study of diopsides in abyssal peridotites. *J. Geophys. Res.* **95**, 2661–2678 (1990).
26. Rampone, E. *et al.* Trace element and isotope geochemistry of depleted peridotites from an N-MORB type ophiolite (Internal Liguride, N. Italy). *Contrib. Mineral. Petrol.* **123**, 61–76 (1996).
27. Brenan, J. M., Shaw, H. F. & Ryerson, F. J. in *7th Annual Goldschmidt Conference* 36 (Tucson, AZ, USA, 1997).
28. Schneider, M. E. & Eggler, D. H. Fluids in equilibrium with peridotite minerals: implications for mantle metasomatism. *Geochim. Cosmochim. Acta* **50**, 711–724 (1986).
29. Keppler, H. Constraints from partitioning experiments on the composition of subduction-zone fluids. *Nature* **380**, 237–240 (1996).
30. Ryan, J. G., Morris, J., Tera, F., Leeman, W. P. & Tsvetkov, A. Cross-arc geochemical variations in the Kurile arc as a function of slab depth. *Science* **270**, 625–627 (1995).
31. Draper, D. S. & Green, T. H. P – T phase relations of silicic, alkaline, aluminous mantle-xenolith glasses under anhydrous and C–O–H fluid-saturated conditions. *J. Petrol.* **38**, 1187–1224.
32. Brenan, J. M., Shaw, H. F., Ryerson, F. J. & Phinney, D. L. Mineral–aqueous fluid partitioning of trace elements at 900 °C and 2.0 GPa: constraints on the trace element chemistry of mantle and deep crustal fluids. *Geochim. Cosmochim. Acta* **59**, 3331–3350 (1995).
33. Hawkesworth, C. J., Gallagher, K., Hergt, J. M. & McDermott, F. Destructive plate margin magmatism: geochemistry and melt generation. *Lithos* **33**, 169–188 (1994).
34. Eiler, J. M., Graham, C. M. & Valley, J. W. SIMS analysis of oxygen isotopes: matrix effects in complex minerals and glasses. *Chem. Geol.* **138**, 221–244 (1997).
35. McDonough, W. F. & Sun, S. S. The composition of the Earth. *Chem. Geol.* **120**, 223–253 (1995).
36. Plank, T. & Langmuir, C. The chemical composition of subducting sediment and its consequences for the crust and mantle. *Chem. Geol.* (in the press).

Correspondence and requests for materials should be addressed to J.M.E. (e-mail: eiler@mail.gps.caltech.edu).

Skull of a Jurassic ankylosaur (Dinosauria)

Kenneth Carpenter, Clifford Miles* & Karen Cloward

Department of Earth Sciences, Denver Museum of Natural History,
2001 Colorado Boulevard, Denver, Colorado 80205, USA

* Western Paleontological Laboratories, 1038 North Industrial Park Drive,
Orem, Utah 84057, USA

The origin and early evolution of many major dinosaur groups are poorly known because specimens are rare. One of these groups, the Ankylosauria, or armour-plated dinosaurs, is best known from well-preserved specimens from the Upper Cretaceous period of Asia and North America. Here we describe a well-preserved skull of an earlier, Late Jurassic ankylosaur, which will be important in clarifying the early history of this group. The specimen, *Gargyleosaurus parkpini* gen. et sp. nov., was collected from the Upper Jurassic Morrison Formation of Wyoming, USA. Despite its geological age, the skull shows features seen in Late Cretaceous ankylosaurs, including fusion of bone armour to the surface of the skull and mandible and closure of two skull openings, the antorbital and upper temporal fenestrae. The new taxon also has characters common to the two ankylosaur families, the Ankylosauridae and Nodosauridae¹⁻⁴, supporting the proposal that the Ankylosauria originated from a single ancestor²⁻⁴. Nevertheless, specialized characters place *Gargyleosaurus* as the most primitive, or basal, member of the Ankylosauridae.

The two families of Ankylosauria^{1,2} are thought, from study of fragmentary remains from China and Europe^{5,6}, to have evolved from a thyreophoran ancestor during the Middle Jurassic period. The scrappiness of these remains adds little to our understanding of these early ankylosaurs and it is not certain that the material has been correctly assigned to either family. Here, however, we describe a new species, *Gargyleosaurus*, on the basis of a complete skull, the first known for any Jurassic ankylosaur, and a significant portion of the postcranial skeleton. Despite the small size of the specimen

(~3 m long), the closure of the sutures in the vertebral column indicates that the individual had reached adult size.

Ornithischia
Ankylosauria
Ankylosauridae

Gargyleosaurus parkpini gen. et sp. nov.

Etymology. *Gargyleosaurus* (Latin): the 'gargoyle reptile', in reference to the gargoyle-like appearance of the skull in profile; *parkpini* (Latin): for J. Parker and T. Pinegar who discovered the holotype.

Holotype. DMNH 27726 (Denver Museum of Natural History, Denver, CO), a partial skeleton with skull articulated with the first three cervical vertebrae and first two cervical rings.

Horizon and locality. Upper part of the Morrison Formation, Bone Cabin Quarry, Albany County Wyoming, USA. Associated dinosaur fauna includes typical Morrison forms, such as *Apatosaurus*, *Brachiosaurus*, *Camarasaurus*, *Diplodocus*, *Allosaurus*, *Ornitholestes*, *Dryosaurus* and *Stegosaurus*.

Diagnosis. *Gargyleosaurus* differs from all other ankylosaurids in the presence of seven premaxillary teeth; the crown base of the cheek teeth are not swollen into the typical ankylosaur cingulum; the premaxilla has a notch in the shape of an inverted U at the front; there is no secondary palate in the roof of the mouth; nostril openings face laterally as in *Ankylosaurus*, *Shamosaurus* and most nodosaurids; the vomer is not expanded upwardly to bisect the nasal cavity; the air passage is straight rather than convoluted as seen in later ankylosaurids; centra of all vertebrae are elongated, unlike the short, wide centra of *Mymoorapelta*; tips of caudal neural spines are not expanded as in *Mymoorapelta*⁷; the first armour ring of the neck consists of five partly fused plates; the second armour ring consists of six partly fused plates; at least two elongated spines extend from each shoulder; body armour consists of thin walled cones, unlike the solid conical armour of *Mymoorapelta*.

Gargyleosaurus shows a mixture of characters seen in the later Cretaceous families of ankylosaurs, the Nodosauridae and Ankylosauridae. From the top, the skull resembles that of many Late Cretaceous ankylosaurids¹ in its triangular shape and in the presence of large, triangular scutes at the rear corners. The cheek tooth row is inset very deeply, as also occurs in ankylosaurids (Fig. 1). However, the simple leaf-shaped cheek teeth with slightly expanded

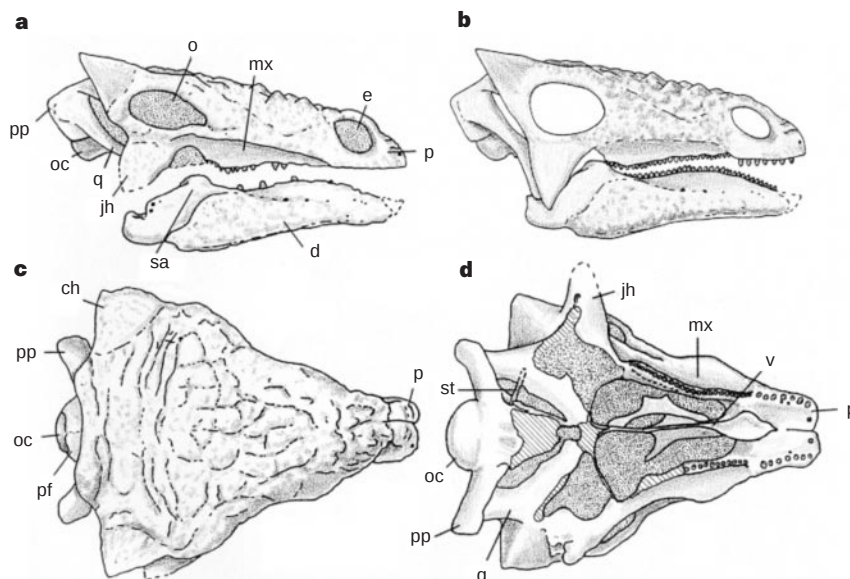


Figure 1 Skull and lower jaw of *Gargyleosaurus parkpini*, specimen DMNH 27726. **a**, Right side as preserved; **b**, right side restored; **c**, skull in top view; **d**, skull in ventral view. Scale, 10 cm. Damaged areas are cross-hatched and matrix is

stippled. Abbreviations: ch, cranial 'horn'; d, dentary; e, external nares; jh, jugal 'horn'; mx, maxilla; o, orbit; oc, occipital condyle; p, premaxilla; pf, proatlas facet; pp, paroccipital process; q, quadrate; sa, surangular; st, stapes; v, vomer.

Table 1 Distribution of characters used in the phyletic analysis

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
<i>Scelidosaurus</i>	0	0	0	0	0	0	0	?	0	0	0	0	0	?	0	0	?	?	0	0	0	0	0	0	0	0
<i>Gargoylesaurus</i>	1	1	1	1	1	0	1	0	0	0	1	1	1	0	0	1	0	0	0	1	0	1	0	0	0	1
Nodosauridae	0	0	1	1	1	0	0	1	1	0	1	0	1	0	1	1	0	1	0	1	1	0	0	0	1	0
Ankylosauridae	1	1	1	0	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1

Characters 1–26 are defined in the Methods.

bases are reminiscent of those found in primitive ornithischians, such as *Lesothosaurus* from the Upper Triassic of Africa; in later ankylosaurids, the crowns of the teeth are much more bulbous or swollen, forming a cingulum. Furthermore, the conical, slightly compressed teeth of the premaxilla have denticles along their margins, making these teeth similar in appearance to those in the primitive stegosaur *Huayangosaurus* (such teeth are also seen in the rear-most premaxillary teeth of *Lesothosaurus*). This similarity lends supports to the proposed relationship between the Stegosauria and Ankylosauria^{2,3}.

The scooped beak formed by the premaxillary bones in *Gargoylesaurus* is long and narrow. Such a beak is also seen in Cretaceous nodosaurids; Late Cretaceous ankylosaurids, on the other hand, typically have a short, wide premaxillary scoop. Within the snout, the complex, looped air passages formed by the secondary palate seen in some (but not all) Cretaceous ankylosaurids is not present. Computerized axial tomography (CAT) scan reveals that the secondary palate did not develop and that the air passage is simple and direct. In addition, the vomer plate, formed by a vertical sheet of bone in the palate, does not extend to the underside of the skull roof as it does in nodosaurids (except *Pawpawsaurus*⁸) and in ankylosaurids¹. Instead, it forms a simple 'T' with the palatines as it does in less specialized ornithischians, such as *Hypsilophodon*⁹. CAT scans also show that the brain had a sharp bend or flexure similar to that seen in *Hypsilophodon*¹⁰ and the nodosaurids *Struthiosaurus*¹¹ and *Polacanthus*¹², rather than the moderate flexure seen in the ankylosaurid *Euoplocephalus*¹³.

The discovery of *Gargoylesaurus* and its contemporary *Mymoorapelta* shows that ankylosaur evolution was well underway by the Late Jurassic period. The nodosaurid-like features in *Gargoylesaurus*, such as premaxillary teeth and sharp flexure of the brain, are apparently primitive for the Ankylosauria because they also occur in other ornithischian dinosaurs. The triangular-shaped head, large triangular armour on the rear corners of the skull and deeply inset cheek teeth indicate a close relationship between *Gargoylesaurus* and Late Cretaceous ankylosaurids. The next step in ankylosaur evolution, the radiation of the two families in the

Cretaceous, is now being investigated in a stratigraphic succession of ankylosaurids in the Lower Cretaceous Cedar Mountain Formation of eastern Utah^{14–16}. At least one of these ankylosaurids shows close affinities with *Gargoylesaurus*, although a complete analysis has yet to be finished (J. Kirkland, manuscript in preparation). □

Methods

We determined the phylogenetic position of *Gargoylesaurus* from an analysis of 26 cranial characters using Hennig86. We created a single robust cladistic tree with *Gargoylesaurus* positioned at the base of the family Ankylosauridae (Fig. 2). Our analysis (see Table 1) did not include postcranial characters because much of the specimen is still being prepared. Nevertheless, analysis using these characters is not expected to change the results because so many of the characters separating the two ankylosaur families are cranial¹². 1, Skull width equal to or greater than 3/4 length; 2, two pairs of plates on the nasal region; 3, large horns at top-rear corners; 4, notch on skull roof for lateral temporal fenestra; 5, large horn on jugal; 6, lateral temporal fenestra closed; 7, pattern of cranial armour asymmetrical; 8, smooth armour on skull roof, posterior to orbits; 9, large, central cranial plate; 10, external nares face anterior; 11, paroccipital project posterolaterally; 12, posterior overhang of occiput; 13, quadrate fused to paroccipital process; 14, beak edge not continuous with maxillary tooth row; 15, premaxillary teeth absent; 16, deeply inset cheek tooth rows; 17, vomer plate (sagittal septum) in palate contacts underside of snout roof; 18, secondary palate; 19, sinus developed in the maxilla; 20, quadrate sloped anteriorly; 21, tooth crowns swollen or with a cingulum; 22, occipital condyle without neck; 23, snout arched in profile view; 24, coronoid process of lower jaws low; 25, maxillary tooth rows give palate an hourglass shape; 26, occipital condyle hemispherical.

Received 29 January; accepted 28 March 1998

- Coombs, W. & Maryanska, T. in *The Dinosauria* (eds Weishampel, D., Dodson, P. & Osmolska, H.) 456–483 (Univ. California Press, Berkeley, 1990).
- Coombs, W. The families of the ornithischian dinosaur order Ankylosauria. *Palaeontology* **21**, 143–170 (1978).
- Norman, D. B. in *Third Symp. on Mesozoic Terrest. Ecosyst., Short Papers* (eds Reif, W. E. & Westphal, F.) 157–162 (Attempto, Tübingen, 1984).
- Sereno, P. C. Phylogeny of the bird-hipped dinosaurs (Order Ornithischia). *Nat. Geogr. Res.* **2**, 234–256 (1986).
- Dong, S. An ankylosaur (ornithischian dinosaur) from the Middle Jurassic of the Junggar Basin, China. *Vert. Pal. Asiatica* **10**, 257–266 (1993).
- Galton, P. Armoured dinosaurs (Ornithischia: Ankylosauria) from the Middle and Upper Jurassic of Europe. *Palaeontol. Abt. A* **182**, 1–25 (1983).
- Kirkland, J., Carpenter, K., Hupps, K., Hunt, A. & Scheetz, R. Ankylosaurs from the Morrison Formation. *Mod. Geol.* (in the press).
- Lee, Y.-M. A new nodosaurid ankylosaur (Dinosauria: Ornithischia) from the Paw Paw Formation (Late Albian) of Texas. *J. Vert. Paleontol.* **16**, 232–245 (1996).
- Galton, P. The Ornithischian dinosaur *Hypsilophodon* from the Wealden of the Isle of Wight. *Bull. Br. Mus. Nat. Hist.* **25**, 1–152 (1974).
- Galton, P. Cranial and endocranial casts from ornithomimid dinosaurs of the families Dryosauridae and Hypsilophodontidae (Reptilia: Ornithischia). *Geol. Palaeontol.* **23**, 217–239 (1989).
- Pereda-Suberbiola, J. & Galton, P. M. A revision of the cranial features of the dinosaur *Struthiosaurus austriacus* Bunzel (Ornithischia: Ankylosauria) from the Late Cretaceous of Europe. *Neues Jb. Geol. Palaeontol. Mh.* **191**, 173–200 (1994).
- Norman, D. B. & Faiers, T. On the first partial skull of an ankylosaurian dinosaur from the Lower Cretaceous of the Isle of Wight, southern England. *Geol. Mag.* **133**, 299–310 (1996).
- Coombs, W. An endocranial cast of *Euoplocephalus* (Reptilia, Ornithischia). *Palaeontol. Abt. A* **161**, 176–182 (1978).
- Kirkland, J., Carpenter, K. & Burge, D. A nodosaur with a distinct sacral shield of fused armor from the Lower Cretaceous of east-central Utah. *J. Vert. Paleontol.* **11** (suppl. 3), 40A (1991).
- Carpenter, K., Kirkland, J., Miles, C., Cloward, K. & Burge, D. Evolutionary significance of new ankylosaurs (Dinosauria) from the Upper Jurassic and Lower Cretaceous, Western Interior. *J. Vert. Paleontol.* **19** (suppl. 3), 25A (1996).
- Carpenter, K., Kirkland, J., Burge, D. & Bird, J. in *Vertebrate Fossils of Utah* (ed. Gillette, D.) (Utah Geological Survey, in the press).

Acknowledgements. We thank J. Kirkland for discussions about early ankylosaurs, and D. Norman for comments on the manuscript. K.C. thanks the Western Paleontological Labs for recognizing the scientific importance of the specimen and for donating it to the Denver Museum of Natural History.

Correspondence and requests for materials should be addressed to K.C. (e-mail: crpntr@ix.netcom.com).

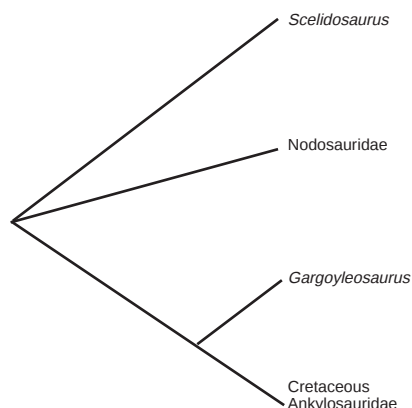


Figure 2 Cladogram showing the phylogenetic position of *Gargoylesaurus* on the basis of 26 cranial characters (most of the postcranial skeleton is still in preparation).

Comparative evidence for the evolution of genitalia by sexual selection

Göran Arnqvist

Department of Animal Ecology, University of Umeå, S-901 87 Umeå, Sweden

Rapid divergent evolution of male genitalia is one of the most general evolutionary trends in animals with internal fertilization; the shapes of genital traits often provide the only reliable characters for species identification¹. Yet the evolutionary processes responsible for this pattern remain obscure. The long-standing lock-and-key hypothesis, still popular among taxonomists, suggests that genitalia evolve by pre-insemination hybridization avoidance; that is, hybrid inferiority drives the evolution of male genitalia with a proper mechanical fit to female genitalia. The sexual selection hypothesis^{2,3}, in contrast, proposes that divergent evolution of genitalia is the result of sexual selection, brought about by variation in postinsemination paternity success among males. Here, by comparing pairs of related clades of insects that differ in mating system, I assess how the opportunity for postmating sexual selection affects the rate of divergent evolution of male genitalia. Genital evolution is more than twice as divergent in groups in which females mate several times than in groups in which females mate only once. This pattern is not found for other morphological traits. These findings provide strong empirical evidence in favour of a postmating sexual selection mechanism of genital evolution.

Under the postmating sexual selection hypothesis, selection on male genitalia is caused by mechanisms that generate variation in postinsemination paternity success among males. Such mechanisms include: first, any of several female processes that affect male paternity success (that is, cryptic female choice³⁻⁵); second, com-

petition between male gametes for fertilization (that is, sperm competition^{6,7}); and third, evolutionary arms races between males and females over the control of fertilization (that is, sexual conflict^{4,8-10}). The key prediction of this hypothesis concerns the relationship between mating system and the rate of genital evolution^{1,3}. In taxa in which females typically mate with only one male (monandry), there can be little variation in male postinsemination paternity success and postmating sexual selection on genitalia will thus be weak or absent. If females mate with many males (polyandry), on the other hand, there will be ample opportunity for variation in male postinsemination paternity success and therefore for postmating sexual selection also. Under the lock-and-key hypothesis, selection for hybridization avoidance is suggested to impel the evolution of male genitalia with a proper mechanical fit. In contrast to postmating sexual selection, such selection for pre-insemination reproductive isolation would be expected to be more intense in monandrous species than in polyandrous species. A given occurrence of interspecific matings will generally be more evenly distributed among polyandrous females than among monandrous females, leading to lower variation in female fitness in polyandrous species and therefore to a weaker selection for pre-insemination reproductive isolation. Here I analyse a series of phylogenetic contrast, comparing morphological divergence in pairs of related clades of insects with differing mating systems (Fig. 1a). This is the first general quantitative assessment of the rate of genital evolution under polyandry relative to that under monandry.

Comparisons of the rate of evolutionary divergence of complex morphological traits in a set of related species have been hampered by problems with identifying homologous structures, as well as by a lack of appropriate methods for quantifying shape variation. Previous comparative studies have often resorted to various subjective ratings of morphological complexity¹¹⁻¹³. Here I use one of the new tools of geometric morphometrics¹⁴, which not only provides objective and quantitative descriptors of shape but also avoids the problem of defining homologous landmarks (that is, structural points with correspondence resulting from descent from the same point in a common ancestor) across species^{14,15}. By

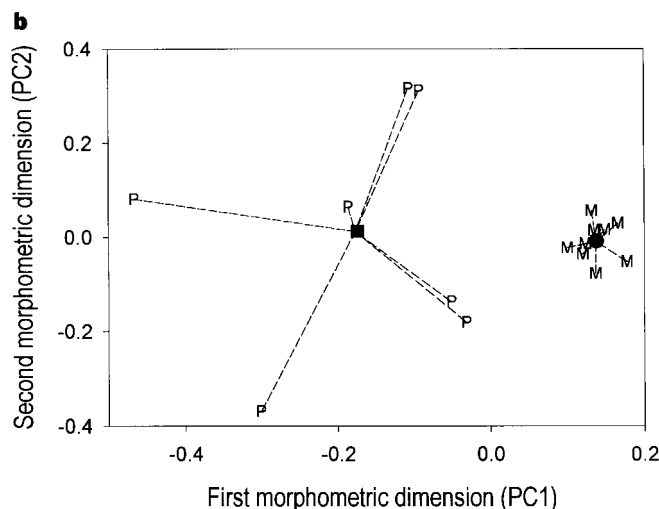
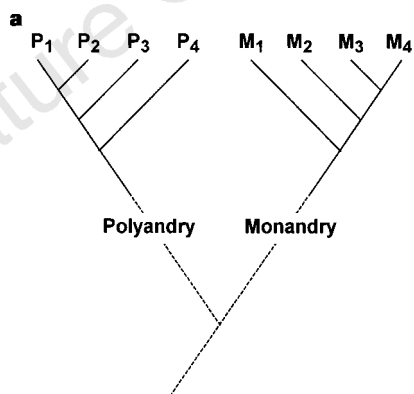


Figure 1 Comparison of the rate of genitalic evolution in polyandrous and monandrous clades. **a**, This study is based on a number of phylogenetic contrasts, where pairs of clades that share a common ancestry are compared. In each contrast, a measure of interspecific morphological dissimilarity within a clade that exhibits a polyandrous mating system (P_{1-4}) is divided by the same measure in a related monandrous clade (M_{1-4}) to form a morphometric distance ratio. Thus, dissimilarities between the two clades are ignored. A ratio higher than unity implies that morphological evolution has been more divergent in the polyandrous clade. **b**, Morphological dissimilarity within a clade was measured

as the average Euclidean distance from the species to the mean of the clade in a common multidimensional shape space. The procedure is illustrated here, in two dimensions only, for male genitalia of several species in the two Dipteran genera *Dryomyza* (polyandrous, P) and *Lucilia* (monandrous, M). Dashed lines represent the distances from each species to the mean (filled symbols) of the clade. In this case, species in the polyandrous clade are about four times as different from one another as are the species in the monandrous clade (see Table 1). PC, principal component.

Table 1 Morphometric distance ratios for genital and general traits for 19 different phylogenetic contrasts

Polyandrous clade			Monandrous clade		Morphometric distance ratio	
Order	Family	Genus	Family	Genus	Genital trait	Other trait
Ephemeroptera	Siphonuridae	<i>Siphonurus</i> (4)	Caenidae	<i>Caneis</i> (4)	1.51	1.13
Lepidoptera	Pieridae	<i>Pieris</i> (3)	Satyridae	<i>Lasiommata</i> (3)	17.60	0.07
	Pieridae	<i>Colias</i> (4)	Satyridae	<i>Coenonympha</i> (4)	3.23	0.08
	Nymphalidae	<i>Eueides</i> (4)	Nymphalidae	<i>Eueides</i> (4)	1.47	1.42
	Nymphalidae	<i>Heliconius</i> (14)	Nymphalidae	<i>Heliconius</i> (13)	1.34	0.80
	Tortricidae	<i>Choristoneura</i> (4)	Tortricidae	<i>Epiphyas</i> (4)	1.50	—*
	Noctuidae	<i>Euoxa</i> (5)	Psychidae	<i>Deborrea</i> (5)	1.23	0.50
	Noctuidae	<i>Helicoverpa</i> (8)	Psychidae	<i>Cryptothoele</i> (7)	1.02	0.83
	Noctuidae	<i>Heliothis</i> (9)	Psychidae	<i>Oiketicus</i> (9)	1.21	0.49
Diptera	Dryomyzidae	<i>Dryomyza</i> (7)	Calliphoridae	<i>Lucilia</i> (9)	4.26	—*
	Drosophilidae	<i>Drosophila</i> (23)	Culicidae	<i>Anopheles</i> (28)	1.52	—*
	Tephritidae	<i>Rhagoletis</i> (24)	Tephritidae	<i>Bactrocera</i> (9)	1.31	1.27
	Anthomyiidae	<i>Coenosia</i> (10)	Anthomyiidae	<i>Delia</i> (11)	3.80	—*
	Cecidomyiidae	<i>Rhopalomyia</i> (13)	Cecidomyiidae	<i>Mayetiola</i> (7)	1.53	—*
	Chironomidae	<i>Stictochironomus</i> (6)	Chironomidae	<i>Clunio</i> (6)	0.93	0.49
	Chironomidae	<i>Chironomus</i> (17)	Chironomidae	<i>Pontomyia</i> (4)	2.08	—*
Coleoptera	Anobiidae	<i>Ernobius</i> (16)	Anobiidae	<i>Xestobium</i> (3)	4.34	30.29
	Elateridae	<i>Agriotes</i> (6)	Elateridae	<i>Ctenicera</i> (5)	1.39	—*
	Dermestidae	<i>Dermestes</i> (4)	Dermestidae	<i>Trogoderma</i> (9)	1.31	0.46
Average morphometric ratio (R):					2.19	0.72
Non-parametric tests of $H_0: R = 1$					$P < 0.001$	$P = 0.388$
Parametric tests of $H_0: R = 1$					$P < 0.001$	$P = 0.424$

The morphometric distance ratio represents the average Euclidean distance of the species in a polyandrous clade to their mean (centroid), divided by the corresponding average distance for the monandrous clade in the contrast. Non-parametric tests of average morphometric ratios were performed with sign tests, and parametric tests with t -tests of $\log_{10}$ transformed data. Numbers within parentheses represent the numbers of species in each genus included in the analysis. A list of all species included, and a specification of the traits used, can be found as Supplementary information or obtained from the author on request.

* No comparable data available.

describing the outlines of the genitalia of each species with a nonlinear function (see Methods), and by subsequently analysing morphological shape variation among species as variance in the parameters of the fitted functions, this method allows the ordination of all the species in each contrast in a common multivariate morphological shape space (Fig. 1b).

The results of this analysis show that male genitalia evolve much more divergently in taxa in which females mate many times. The shape of male genitalia of polyandrous species were more dissimilar than were those of monandrous species in 18 out of 19 contrasts, and the average morphological distance between the genitalia of polyandrous species was more than twice that of monandrous species (see Table 1 for tests). This pattern did not differ between orders (Kruskal–Wallis analysis of variance, $P = 0.84$), and the taxonomic distance between the two clades in each contrast did not significantly affect the relative degree of genital divergence within clades (within versus between-family contrasts; Mann–Whitney U -test, $P = 0.80$). There was no association between the distance ratios of genitalia and the distance ratios of other traits across contrasts (Spearman rank correlation, $P > 0.9$). The analysis did not reveal any influence of mating system on evolutionary divergence for morphological traits other than genital traits (Table 1), and the distance ratios of genital traits were indeed significantly larger than those of other traits (paired Wilcoxon signed rank test, $P = 0.023$; Kolmogorov–Smirnov two-sample test, $P = 0.003$).

Many factors other than selection could potentially influence measures of interspecific evolutionary divergence within a given clade (such as age of clade, biogeographic characteristics, taxonomic resolution, genetic architecture and mating-system characteristics other than female mating frequency). Given the confounding role of such factors, it is remarkable that monandrous and polyandrous clades differed so consistently in the relative rates of divergent evolution of male genitalia. This study offers three important and consequential insights. First, it provides strong evidence in favour of the postmating sexual selection mechanism of genital evolution, and thus enhances our knowledge of the processes behind this general evolutionary trend^{1–3,16}. Future research should attempt to determine which forms of postmating sexual selection are responsible for genital evolution^{3,10,16}. Second, the results indicate that

the same process (that is, sexual selection) may be responsible for the evolutionary elaboration of both primary and secondary sexual traits, suggesting that this old dichotomy, which Darwin¹⁷ realized was problematic but nevertheless adopted, should be reconsidered^{16,18}. Third, as traits evolving by sexual selection tend to be more phenotypically and genetically variable than other traits^{19,20}, this study calls for quantifications of the degree of intraspecific variability in genital traits. The prevailing typological view of intromittent genitalia in taxonomy, especially in species definitions, may need to be reconsidered. □

Methods

Case selection and data acquisition. I searched for clades suitable for the phylogenetic contrasts^{21,22} in previous reviews²³ and comparative studies^{24–26}, as well as in reference databases and on the internet. Three criteria had to be met for inclusion of a clade in the analysis. First, reliable data on female mating frequencies had to be at hand, typically in the form of female spermatophore/ejaculate counts in natural populations or detailed field and/or laboratory studies of mating behaviour. Second, the phylogeny of the species included in a given contrast had to be well established. Third, taxonomic revisions containing high-quality illustrations of male genitalia had to be existent, as such illustrations were used to characterize each species.

I located 19 phylogenetic contrasts, representing four different orders, that met these criteria; all of these contrasts were independent in the sense that no clade was represented in more than one contrast (Table 1). This selection was based on a large number of published articles, as well as on personal contacts with a large number of colleagues. A complete list of these sources can be found as Supplementary information or obtained from the author on request. For each species, I captured the outlines of two trait types (male genitalia and, when available, a general trait) with a digitizing tablet (Summasketch III), using illustrations presented in published taxonomic revisions. For consistency, only one such source was used for each clade to avoid artifactual intraclade variation in morphology. The general traits were wings (nine contrasts), body parts (two contrasts) or legs (one contrast). When more than one general trait was at hand, the trait that exhibited most interspecific divergence between species in the contrast was included in the analysis.

Elliptic Fourier analysis. For each contrast and trait type, the outlines of all species were included in a common elliptic Fourier analysis^{14,15,27}, using the software EFA-Win (<http://life.bio.sunysb.edu/morph/soft-out.html>). The

Fourier analyses were made invariant of size, position and rotation, and all used 30 harmonics (yielding 120 Fourier coefficients). These functions provided a near perfect fit even to the most complex outlines.

Multivariate ordination. For each contrast and trait type, the 120 Fourier coefficients for each species were treated as variables in a principal component analysis, performed on the covariance matrix^{28,29}. The first seven principal components, collectively describing >98% of the shape variation in all cases, were retained for ordination. These principal components form orthogonal dimensions in a multidimensional shape space, where all species occupy a given location (Fig. 1b). To quantify the morphological dissimilarity between the polyandrous species relative to that of the monandrous species, I calculated a morphometric distance ratio, representing the average Euclidean distance in this multidimensional space of the species in the polyandrous clade to their mean (centroid) divided by the corresponding average distance for the monandrous clade in the contrast (Fig. 1b). This ratio measures the amount of morphological variance within the polyandrous clade relative to that within the monandrous clade, ignoring variance due to differences between the clades. In contrasts in which the number of species in the two clades were skewed (difference > 1), the elliptic Fourier analysis and the subsequent multivariate ordination procedure were repeated many times (>10) using a random subsample from the more speciose clade to match the number of species in the less speciose clade, and the average morphometric distance ratio from these repeated measures was used for analysis.

Received 29 January; accepted 23 March 1998.

1. Eberhard, W. G. *Sexual Selection and Animal Genitalia* (Harvard Univ. Press, Cambridge, MA, 1985).
2. Eberhard, W. G. Evaluating models of sexual selection: genitalia as a test case. *Am. Nat.* **142**, 564–571 (1993).
3. Eberhard, W. G. *Females Control: Sexual Selection by Cryptic Female Choice* (Princeton Univ. Press, Princeton, NJ, 1996).
4. Lloyd, J. E. Mating behavior and natural selection. *Fla. Entomol.* **62**, 17–34 (1979).
5. Thornhill, R. Cryptic female choice and its implications in the scorpionfly *Harporhynchus nigricaps*. *Am. Nat.* **122**, 765–788 (1983).
6. Waage, J. K. Dual function of the damselfly penis: sperm removal and transfer. *Science* **203**, 916–918 (1979).
7. Gage, M. J. Removal of rival sperm during copulation in a beetle, *Tenebrio molitor*. *Anim. Behav.* **44**, 587–589 (1992).
8. Arnqvist, G. & Rowe, L. Sexual conflict and arms races between the sexes: a morphological adaptation for control of mating in a female insect. *Proc. R. Soc. Lond. B* **261**, 123–127 (1995).
9. Rice, W. R. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nature* **381**, 232–234 (1996).
10. Alexander, R. D., Marshall, D. C. & Cooley, J. R. in *The Evolution of Mating Systems in Insects and Arachnids* (eds Choe, J. C. & Crespi, B. J.) 4–31 (Cambridge Univ. Press 1997).
11. Verell, P. A. Primate penile morphologies and social systems: further evidence for an association. *Folia Primatol. (Basel)* **59**, 114–120 (1992).
12. Harcourt, A. H. & Gardiner, J. Sexual selection and genital anatomy of male primates. *Proc. R. Soc. Lond. B* **255**, 47–53 (1994).
13. Proctor, H. C., Baker, R. L. & Gwynne, D. T. Mating behaviour and spermatophore morphology: a comparative test of the female-choice hypothesis. *Can. J. Zool.* **73**, 2010–2020 (1995).
14. Lestrel, P. E. (ed) *Fourier Descriptors and their Applications in Biology* (Cambridge Univ. Press, 1997).
15. Rohlf, F. J. in *Ordination in the Study of Morphology, Evolution and Systematics of Insects* (eds Sørensen, J. T. & Foottit, R.) 95–112 (Elsevier, Amsterdam, 1992).
16. Arnqvist, G. The evolution of animal genitalia: distinguishing between hypotheses by single species studies. *Biol. J. Linn. Soc.* **60**, 365–379 (1997).
17. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (Murray, London, 1871).
18. Andersson, M. *Sexual Selection* (Princeton Univ. Press, Princeton, NJ, 1994).
19. Pomiankowski, A. & Møller, A. P. A resolution of the lek paradox. *Proc. R. Soc. Lond. B* **260**, 21–29 (1995).
20. Rowe, L. & Houle, D. The lek paradox and the capture of genetic variance by condition dependent traits. *Proc. R. Soc. Lond. B* **263**, 1415–1421 (1996).
21. Burt, A. Comparative methods using phylogenetically independent contrasts. *Oxf. Surv. Evol. Biol.* **6**, 33–53 (1989).
22. Harvey, P. H. & Pagel, M. D. *The Comparative Method in Evolutionary Biology* (Oxford Univ. Press, 1991).
23. Thornhill, R. & Alcock, J. *The Evolution of Insect Mating Systems* (Harvard Univ. Press, Cambridge, MA, 1983).
24. Ridley, M. The timing and frequency of mating in insects. *Anim. Behav.* **37**, 535–545 (1989).
25. Ridley, M. The incidence of sperm displacement in insects: four conjectures, one corroboration. *Biol. J. Linn. Soc.* **38**, 349–367 (1989).
26. Ridley, M. The control and frequency of mating in insects. *Funct. Ecol.* **4**, 75–84 (1990).
27. Person, S., Rohlf, F. J. & Koehn, R. K. Measuring shape variation of two-dimensional outlines. *Syst. Zool.* **34**, 59–68 (1985).
28. Rohlf, F. J. & Archie, J. A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Syst. Zool.* **33**, 302–317 (1984).
29. Liu, J. et al. Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila similans* and *D. mauritiana*. *Genetics* **142**, 1129–1145 (1996).

Supplementary information is available on Nature's World-Wide Website (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

Acknowledgements. I thank L. Rowe and P. Watson for comments. This study was made possible by the expert advice from many entomologists (see Supplementary information), and was generously funded by a grant from the Swedish Natural Science Research Council.

Correspondence and requests for materials should be addressed to the author (e-mail: Goran.Arnqvist@animescol.umu.se).

Smad2 role in mesoderm formation, left-right patterning and craniofacial development

Masatoshi Nomura & En Li

Cardiovascular Research Center, Massachusetts General Hospital East, Department of Medicine, Harvard Medical School, Charlestown, Massachusetts 02129, USA

Signalling by the transforming growth factor- β (TGF- β) superfamily of proteins depends on the phosphorylation and activation of SMAD proteins by heteromeric complexes of ligand-specific type I and type II receptor with serine/threonine-kinase activity¹. The vertebrate SMAD family includes at least nine members, of which Smad2 has been shown to mediate signalling by activin and TGF- β ^{2–5}. In *Xenopus*, Smad2 can induce dorsal mesoderm, mimicking ventral activin and nodal^{2,4}. Here we investigate the function of Smad2 in mammalian development by generating two independent *Smad2* mutant alleles in mice by gene targeting. We show that homozygous mutant embryos fail to form an organized egg cylinder and lack mesoderm, like mutant mice lacking *nodal*^{6,7} or *ActRIB*, the gene encoding the activin type-I receptor⁸. About 20 per cent of *Smad2* heterozygous embryos have severe gastrulation defects and lack mandibles or eyes, indicating that the gene dosage of Smad2 is critical for signalling. Mice *trans*-heterozygous for both *Smad2* and *nodal* mutations display a range of phenotypes, including gastrulation defects, complex craniofacial abnormalities such as cyclopia, and defects in left–right patterning, indicating that Smad2 may mediate nodal signalling in these developmental processes. Our results show that Smad2 function is essential for early development and for several patterning processes in mice.

We generated two independent mutant alleles of *Smad2* by gene targeting in mouse embryonic stem (ES) cells. *Smad2*^{mh1} was produced by replacing part of the conserved MH1 domain with a PGK-neo cassette (Fig. 1a). To test whether this mutation was null, we generated ES cell lines homozygous for *Smad2*^{mh1} (data not shown) and analysed Smad2 expression by western blotting. We found that the amount of Smad2 protein in heterozygous ES cells was significantly reduced. In contrast, neither the normally sized protein nor a truncated form of Smad2 was detected in *Smad2*^{mh1} homozygous ES cells (Fig. 1b), indicating that *Smad2*^{mh1} is a null allele. The second allele, *Smad2*^{mh2-lacZ}, is a 'knock-in' insertion of a *lacZ* reporter gene into the carboxy-terminal MH2 domain created by replacing exon 8 with an *IRES β geo* cassette⁹ (Fig. 2a). The phenotypes resulting from these two mutations were essentially identical.

The offspring from intercrossed *Smad2*^{mh1} heterozygous mice were examined at various developmental stages. Homozygous embryos were recovered at 7.5 days post-coitum (or E7.5), but not at or after E10.5 (Fig. 1c and Table 1), indicating that the

Table 1 Genotype of offspring from the *Smad2*^{mh1} heterozygote crosses

Stage	Total	Resorption	Genotype*		
			+/+	+/-	-/-
E7.5	117	1	33	53 (12)	30
E8.5	39	0	10	21 (4)	8
E10.5–E16.5	203	65	59	79 (9)	0
Newborn	112	–	48	64 (2)	0

* Defects are indicated in parentheses among heterozygous embryos: about 22% (16/74) showed gastrulation defects at E7.5 and E8.5, and about 11% (9/79) showed mandible or eye defects at E10.5–E16.5.

Smad2^{mh1} mutation caused early embryonic lethality. Homozygous mutant embryos at E7.5 and E8.5 were severely retarded and lacked a discernible anterior–posterior axis (Fig. 1d, e). We found that a significant portion (20%) of the *Smad2*^{mh1}/+ embryos and newborn pups showed various abnormalities (Table 1). At E7.5, the affected mutant embryo appeared to grow normally, but the embryonic cylinder was distorted (Fig. 1f). By E8.5, the mutant embryo was severely retarded and deformed, and was often located outside the yolk sac, whereas the extraembryonic tissues including visceral yolk sac were well formed (Fig. 1g). Of the *Smad2*^{mh1}/+ embryos that went through normal gastrulation, about 10% had craniofacial defects, including absence or hypoplasia of mandibles or absence of an eye (Fig. 1h, i).

To exclude the possibility that the defects seen in *Smad2*^{mh1}/+ mice resulted from a dominant-negative effect of a trace amount of a truncated *Smad2* C-terminal domain, we also analysed mice carrying the *Smad2*^{mh2-lacZ} allele (Fig. 2a). We found that the *Smad2*^{mh2-lacZ} mutation resulted in developmental defects in heterozygous and homozygous mutant embryos that were essentially identical to those in *Smad2*^{mh1} mice (Fig. 2b, c), indicating that both *Smad2* mutant alleles are null and the gene dosage of *Smad2* is critical for signalling.

To define the tissue defect, we did histological analysis on embryos from E6.0 to E7.5. At E6.0 and E6.5, the homozygous mutant embryos were considerably smaller than their normal littermates, and the visceral endoderm was detached from the epiblast (Fig. 3a–d). By E7.5, the mesodermal layer, amnion and chorion were formed in wild-type embryos (Fig. 3e), whereas no mesodermal-like cells were present in the mutant, and the epiblast and extraembryonic ectoderm were disorganized (Fig. 3f, g). The proamniotic cavity was initially formed in the mutant embryos

(Fig. 3b) but failed to extend into the extraembryonic ectoderm (Fig. 3d, f, g). In addition, the boundary between epiblast and extraembryonic ectoderm appeared to be disrupted, resulting in ectopic localization of both tissues and the loss of the proximal–distal polarity (Fig. 3f, g). These results indicate that *Smad2* function is required for the appropriate organization of the primitive germ layers in the egg cylinder before gastrulation.

Histological analysis of abnormal E7.5 *Smad2*^{mh1}/+ embryos revealed that mesodermal cells were present in both embryonic and extraembryonic regions: however, the embryonic mesoderm appeared to be unpatterned and the embryo proper was severely retarded (Fig. 3h). In the extraembryonic region, the visceral yolk sac and chorion were normal, whereas amnion and allantois were often not formed (Fig. 3h).

Gastrulation in mice begins at E6.5 with the formation of the primitive streak at the proximal–posterior region of the epiblast (embryonic ectoderm). As gastrulation proceeds, the primitive streak extends anteriorly to the distal tip of the egg cylinder, and mesoderm forms by ingress of epiblast cells through the streak^{10,11}. To see whether the primitive streak was formed in the mutant embryos, we analysed the expression of the early mesodermal marker *Brachyury* (T)¹² by whole-mount *in situ* hybridization. As expected, T expression was confined to the primitive streak in wild-type embryos at E7.5 (Fig. 3i), and later also to the head process and notochord (Fig. 3j). In contrast, the growth-retarded embryos with characteristic *Smad2*^{mh1} homozygous morphology were all negative for T expression (Fig. 3i, j), indicating that the primitive streak was not formed. T expression was detected in *Smad2*^{mh1} heterozygous embryos with gastrulation defects but was restricted to the posterior region of the embryos (Fig. 3k), suggesting that the primitive streak formed initially in the abnormal hetero-

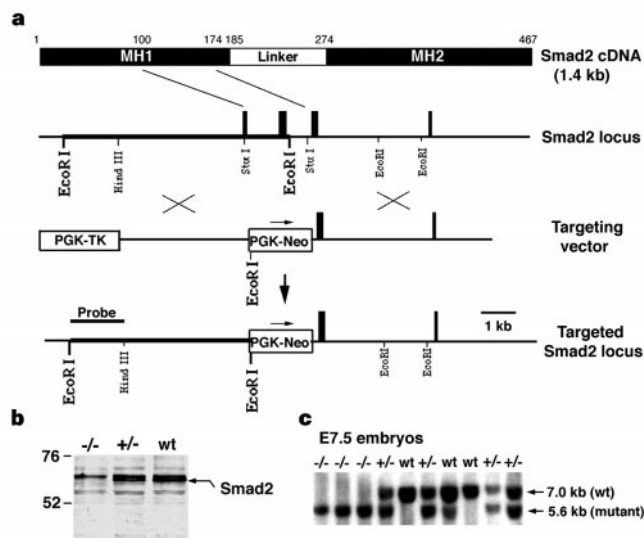
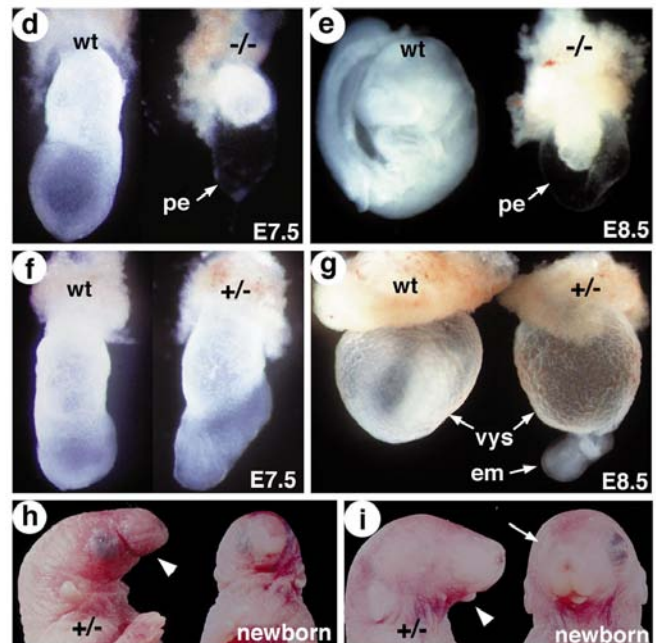


Figure 1 The *Smad2*^{mh1} mutation results in early embryonic lethality and other defects. **a** (top to bottom), *Smad2* cDNA, the wild-type gene locus, the targeting vector, and the *Smad2*^{mh1} mutant allele. An *EcoRI* site was introduced into the targeted *Smad2* locus from the *neo* cassette. The 7.0-kb and 5.6-kb *EcoRI* fragments from the wild-type and the mutant alleles, respectively, are shown as thicker lines. **b**, Western blot analysis of *Smad2* protein in *Smad2*^{mh1} homozygous (–/–), heterozygous (+/–) and wild-type (WT) ES cells. The protein band corresponding to *Smad2* is indicated. Note that neither the full-length nor a truncated *Smad2* protein was detected in the homozygous cells. **c**, Southern blot



analysis of DNA from a litter of E7.5 embryos produced by heterozygote intercrosses. **d, e**, Morphology of *Smad2* homozygous embryos (right) and normal littermates (left) at E7.5 (**d**) and late E8.5 (**e**). **f, g**, Abnormal *Smad2*^{mh1} heterozygous embryos (right) and normal littermates (left) at E7.5 (**f**) and E8.5 (**g**). **h, i**, Eye and mandible defects in *Smad2*^{mh1} heterozygous newborn mice. Arrowheads indicate absence (**h**) or hypoplasia (**i**) of mandibles; arrow indicates the absence of an eye (**i**). em, Embryo; pe, parietal endoderm; vys, visceral yolk sac.

zygous embryos but failed to extend to the distal tip of the embryonic cylinder.

Our results indicate that *Smad2* is required for the organization of the primitive germ layers before gastrulation and for primitive streak formation in mice. Embryos homozygous for *Smad2*^{Robm1}, another mutant *Smad2* allele are able to form extraembryonic mesoderm, visceral yolk sac and chorion¹³, like the *Smad2*^{mh1} or *Smad2*^{mh2-lacZ} heterozygous embryos described here. The apparently less severe defects of *Smad2*^{Robm1} homozygous embryos indicate that *Smad2*^{Robm1} is probably a hypomorph.

Nodal, a member of the TGF- β superfamily¹⁴, has been implicated in primitive streak formation⁶, induction of forebrain¹⁵, and in patterning left–right asymmetry in mice^{16,17}, but the receptor or SMAD responsible for mediating the nodal signal was unknown. The similarity of the early defects evident in mutant *Smad2* homozygotes and in *nodal*-deficient embryos suggests that nodal might signal through *Smad2*: we therefore generated and analysed mice *trans*-heterozygous for both the *Smad2* and *nodal* mutations.

The mutant *nodal*^{lacZ} mice, in which the *nodal* gene was disrupted by insertion of the *IRES* β geo cassette¹⁶, were crossed with *Smad2*^{mh1} mice. Offspring were analysed at various stages from E9.5 to birth. We found that *trans*-heterozygosity of the *Smad2* and *nodal* mutations dramatically enhanced the defects characteristic of *Smad2*^{mh1} mice and caused several new phenotypes to emerge (Fig. 4 and Table 2). At E9.5, 55% (11/20) of the *trans*-heterozygous embryos had gastrulation defects like those seen in abnormal *Smad2*^{mh1} embryos (Fig. 4a). Analysis of *nodal* expression in E7.5 and E8.5 embryos by X-gal staining revealed that the *nodal*^{lacZ} allele was expressed in the posterior region of abnormal *trans*-heterozygous embryos (Fig. 4b, c), probably in the primitive streak region, as judged from the expression pattern of T in abnormal *Smad2*^{mh1} embryos (Fig. 3k). Our finding that the *trans*-heterozygotes and *Smad2*^{mh1} embryos had identical gastrulation defects indicates that *Smad2* and *nodal* may function in the same signalling pathway to regulate gastrulation. Further investigation is needed to determine whether *Smad2* directly mediates nodal signalling.

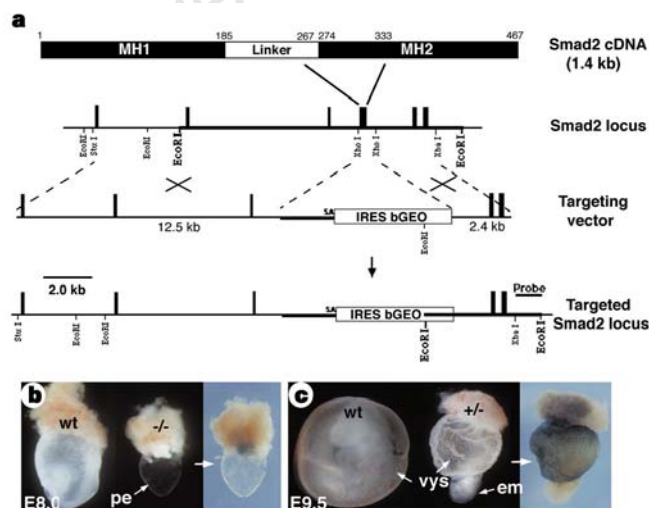


Figure 2 The early developmental defects in *Smad2*^{mh2-lacZ} mutant embryos. **a** (top to bottom), *Smad2* cDNA, the wild-type gene locus, the targeting vector, and the *Smad2*^{mh2-lacZ} mutant allele. An *EcoRI* site was introduced into the targeted *Smad2* locus from the *IRES* β geo cassette. The 14.0-kb and 4.8-kb *EcoRI* fragments from the wild-type and targeted alleles, respectively, are shown as thicker lines. **b, c**, Normal embryos (left) and a homozygous mutant at E8.0 (**b**) or an abnormal heterozygous littermate at E9.5 (**c**) before (middle) or after (right) X-gal staining.

At E15.5–E17.5, 56% (14/25) of the *trans*-heterozygous embryos showed severe craniofacial defects such as cyclopia (9/25) (Fig. 4d); in the most severe cases, the rostral head and eyes were truncated (Fig. 4e). In contrast, only 8% (3/37) of the *Smad2*^{mh1} littermates showed mild mandible defects (Table 2). In addition, some *trans*-heterozygous embryos also had defects in left–right patterning relative to wild-type mice, in which visceral organs are asymmetric in position or shape along the left–right axis: the heart is on the left side of the chest; the right lung has four lobes, whereas the left lung has one; and the stomach is on the left side in the abdomen. Heart looping and embryonic turning are the two early morphogenetic events leading to left–right asymmetry. Of 20 E9.5 *trans*-heterozygous embryos, three displayed abnormal heart looping and opposite turning compared with normal littermates (Fig. 4f). Examination of the visceral organs in E15.5–E17.5 embryos revealed that 32% (8/25) of the *trans*-heterozygotes also had cardiac and laterality defects. The most common cardiac malformation was transposition of the great arteries (TGA) (Fig. 4g), as indicated by

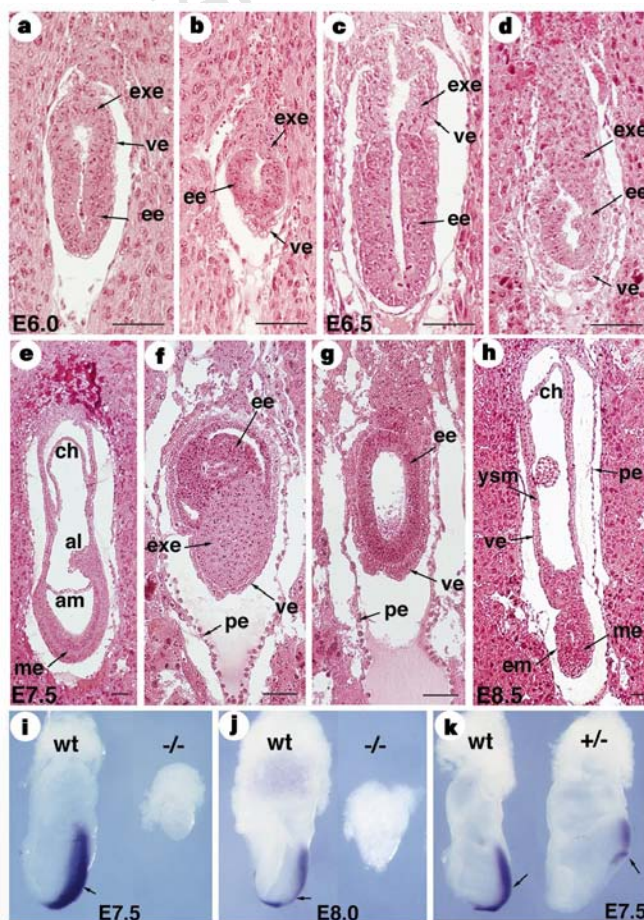


Figure 3 Defects in egg-cylinder organization and primitive streak formation in *Smad2*^{mh1} mutant embryos. **a–g**, Histological analysis of *Smad2*^{mh1} homozygous and control embryos. Sagittal sections of the wild-type (**a, c, e**) and homozygous mutant embryos (**b, d, f, g**) at E6.0 (**a, b**), E6.5 (**c, d**) and E7.5 (**e, f, g**). The epiblast stains slightly darker than extraembryonic ectoderm. In mutant embryos, the embryonic and extraembryonic ectoderm were disorganized (**f, g**). **h**, Histology of an abnormal *Smad2*^{mh1} heterozygous embryo. **i, j**, Whole-mount *in situ* hybridization with the *Brachyury* (T) probe on homozygous embryos at E7.5 (**i**) or E8.0 (**j**). **k**, T expression in an abnormal heterozygous embryo at E7.5 is restricted to the proximal–posterior region of the epiblast. al, Allantois; am, amnion; ch, chorion; ee, embryonic ectoderm; em, embryo; exe, extraembryonic ectoderm; me, mesoderm; pe, parietal endoderm; ve, visceral endoderm; ysm, yolk sac mesoderm. Scale bar, 50 μ m.

the connection of the aorta to the right ventricle and the pulmonary artery to the left ventricle. Six embryos with TGA also had a bilaterally symmetric right lung pattern (right pulmonary isomerism) (Fig. 4h). One of the *trans*-heterozygous embryos had TGA, right pulmonary isomerism, and a right-sided stomach (Fig. 4i). Cyclopia, turning and laterality defects were never seen in *Smad2^{mh1}/+* mice (0/87) or *nodal^{lacZ}/+* mice (0/55) (Tables 1, 2). Cardiac defects such as TGA and right pulmonary isomerism are both characteristic mutant phenotypes of mice lacking the type IIB activin receptor gene¹⁸; similar early gastrulation defects and cranio-

Table 2 Genetic interactions between *nodal^{lacZ}* and *Smad2^{mh1}*

Genotype		Stage		
<i>nodal</i>	<i>Smad2</i>	E9.5	E15.5–17.5	Newborn
WT	WT	22	63	61
+/-	WT	15	55	65
WT	+/-	27 (4)	37 (3)	31
+/-	+/-	20 (14)*	25 (14)†	19
Resorption		2	47	–
Total		86	220	176

The number of abnormal embryos is indicated in parentheses.

* Among the 14 *trans*-heterozygous embryos at E9.5, 11 embryos had gastrulation defects and 3 showed opposite turning.

† At E15.5–E17.5, all 14 abnormal embryos showed severe craniofacial defects. Eight embryos also showed cardiac defects, of which 6 had right pulmonary isomerism in addition to cardiac defects.

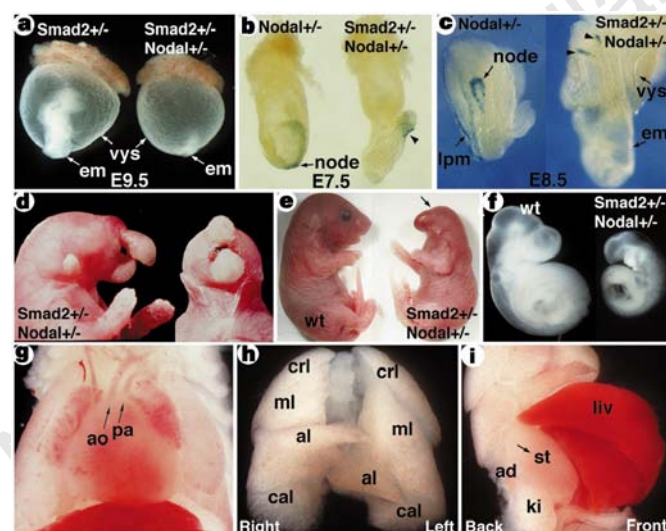


Figure 4 Gastrulation, craniofacial and laterality defects in *Smad2^{mh1}/+, nodal^{lacZ}/+* *trans*-heterozygotes. **a**, An E9.5 *Smad2^{mh1}* heterozygous embryo (left) and a *Smad2^{mh1}/+, nodal^{lacZ}/+* *trans*-heterozygous littermate show similar gastrulation defects. **b**, **c**, Nodal expression in E7.5 (**b**) and E8.5 (**c**) embryos as shown by X-gal staining. LacZ expression was confined to the node (**b**, **c**) and lateral plate mesoderm on the left (**c**) in *nodal^{lacZ}/+* embryos, but was restricted to the posterior region of the *trans*-heterozygous embryos (**b**, **c**, arrowheads). **d**, Side view (left) and front view (right) of a *trans*-heterozygous embryo showing the fusion of two eyes at the midline (cyclopia). **e**, An E17.5 *trans*-heterozygous embryo showing truncation of the anterior head structure, including the eyes (arrow). **f**, An E9.5 *trans*-heterozygous embryo (right) showing opposite turning compared with a normal littermate (tail curving to the opposite side). **g**, An E15.5 *trans*-heterozygous embryo showing transposition of the great arteries. **h**, An E15.5 embryo showing symmetric lobation of left and right lungs. **i**, An E15.5 *trans*-heterozygous embryo showing a right-sided stomach (arrow). ad, Adrenal gland; ao, aorta; em, embryo; ki, kidney; liv, liver; lpm, lateral plate mesoderm; pa, pulmonary artery; st, stomach; vys, visceral yolk sac; cri, ml, cal and al, represent the cranial, medial, caudal and accessory lobes of the right lung, respectively.

facial abnormalities, ranging from an absence of the mandible to cyclopia, as described here, have also been observed in mice with mutations in both type IIA and IIB activin receptors (data not shown). It is possible that Smad2 mediates nodal signalling through type II activin receptors to regulate these developmental processes.

In *Xenopus*, Smad1 mediates signalling by bone morphogenetic protein (BMP) in induction of ventral mesoderm^{2,19}, and Smad2 can induce different mesodermal cell types in a dose-dependent manner like activin and nodal^{4,20,21}. We have provided genetic evidence that Smad2 also functions in a dosage-sensitive manner during mouse development: in the absence of Smad2, formation of the primitive streak and mesoderm is blocked; in heterozygous embryos with half the normal amount of Smad2 (Fig. 1b), the primitive streak formation can be induced. In some embryos, however, probably when the signal transduced by Smad2 is below threshold, primitive streak extension is impaired and epiblast cells differentiate to form predominantly extraembryonic (or ventral type) mesoderm; therefore Smad2 function in mesoderm formation and specification appears to be conserved in vertebrates.

Smad4 (DPC4) is generally believed to be the mediator of signalling by the TGF- β superfamily¹. In *Xenopus*, Smad4 associates with Smad1 in response to BMP and with Smad2 in response to activin in the induction of ventral and dorsal mesoderm, respectively^{22,23}. Targeted mutagenesis of *Smad4* in mice indicates that Smad4 is required initially for the differentiation of visceral endoderm and later for anterior patterning²⁴. However, the early egg-cylinder defect in Smad4-deficient mice appears to be morphologically different from that of Smad2 homozygous embryos, indicating that Smad4 may function earlier in development than Smad2. A more detailed comparison of Smad2 and Smad4 mutant phenotypes should reveal whether Smad4 is also partner of Smad2 during mouse development. □

Methods

Targeting vectors. The mouse *Smad2* genomic DNA was cloned from a 129/Sv genomic library (Stratagene) and the genomic organization determined by Southern blot analysis and DNA sequencing. To generate the *Smad2^{mh1}* targeting vector, a 2.0-kb *StuI*–*StuI* fragment, which contains part of exon 3 and all of exon 4, encoding amino-acid residues 100–174 of Smad2, was deleted and replaced by a PGK-neo-polyA cassette (the *XhoI*–*SalI* fragment from pKJ-1) by blunt-end ligation. The PGK-thymidine kinase gene was inserted into the *HindIII* site 4.0 kb upstream of the *StuI* site. The targeting vector contains 4.0 and 5.5 kb of homologous genomic DNA on either side of the neomycin-resistance cassette. To generate the *Smad2^{mh2-lacZ}* targeting vector, an *XhoI*–*XhoI* fragment, which contains exon 8, encoding amino-acid residues 267–333 of Smad2, was deleted and replaced by an *IRES β geo* cassette (the 6.8-kb *SalI* fragment from pGT1.8es β geo)⁹. The targeting vector contains 12.5 and 2.4 kb of homologous genomic DNA on either side of the *IRES β geo* cassette.

Generation of *Smad2*-disrupted mutant ES cell lines and mice. J1 ES cells were transfected with a linearized targeting vector by electroporation and were selected in medium containing G418 or G418 + FIAU, as described²⁵. For the first targeting vector, 384 G418/FIAU-resistant ES-cell clones were analysed by Southern blot hybridization using the *EcoRI*–*HindIII* genomic fragment as probe (Fig. 1a); one had a targeted allele. For the second targeting vector, of 192 G418-resistant ES-cell clones analysed by Southern blot hybridization using the *XbaI*–*EcoRI* genomic fragment as probe (Fig. 2a), 160 clones showed correct targeting. One *Smad2^{mh1}/+* ES clone and two *Smad2^{mh2-lacZ}/+* ES clones were injected into C57BL/6J blastocysts to obtain chimaeric mice. Germline transmission was achieved with all three clones injected.

Genotype analysis. Genomic DNA was prepared from embryos or tails as described²⁶. For E7.5 embryos, DNA was isolated from whole embryonic and extraembryonic tissues, including parietal yolk sac. For genotype analysis, DNA was digested with *EcoRI*, blotted and hybridized to genomic probes as indicated (Figs 1, 2).

Western blot analysis. Two independent *Smad2^{mh1}* homozygous ES-cell lines were derived from delayed blastocysts (data not shown). ES cells were lysed in TNE buffer (10 mM Tris, pH 7.8, 150 mM NaCl, 1 mM EDTA, 1% N-P40)

containing protease inhibitors. Cell lysates were separated by SDS-PAGE, followed by western blot analysis using a Smad2-specific polyclonal antibody, SED, which was raised against a peptide corresponding to amino-acid residues 227–244 in the linker region²⁷.

Histology and *in situ* hybridization. Embryos were fixed in Bouin's fixative, dehydrated and embedded in paraffin as described²⁸. Serial sections (6 µm thick) were cut and stained with haematoxylin and eosin using standard procedures. For *in situ* hybridization, embryos were dissected out and fixed in 4% paraformaldehyde plus 0.1% Tween 20 in phosphate-buffered saline overnight at 4°C. Strand-specific RNA probes were generated using digoxigenin UTP (Boehringer Mannheim). Whole-mount *in situ* hybridization has been described²⁹.

Generation of *Smad2^{mh1}/+*, *nodal^{lacZ}/+* mutant mice and X-gal staining of embryos. *nodal^{lacZ}/+* mice¹⁶ were crossed with *Smad2^{mh1}/+*. Embryos were dissected at different developmental stages and examined for defects in gastrulation, in craniofacial development or in visceral organs. X-gal staining of E7.5–8.5 embryos has been described¹⁸.

Received 2 March; accepted 22 May 1998.

- Heldin, C. H., Miyazono, K. & ten Dijke, P. TGF-β signalling from cell membrane to nucleus through SMAD proteins. *Nature* **390**, 465–471 (1997).
- Graff, J. M., Bansal, A. & Melton, D. *Xenopus* Mad proteins transduce distinct subsets of signals for the TGFβ superfamily. *Cell* **85**, 479–487 (1996).
- Macias-Silva, M. *et al.* MADR2 is a substrate of the TGFβ receptor and its phosphorylation is required for nuclear accumulation and signaling. *Cell* **87**, 1215–1224 (1996).
- Baker, J. C. & Harland, R. M. A novel mesoderm inducer, *Madr2*, functions in the activin signal transduction pathway. *Genes Dev.* **10**, 1880–1889 (1996).
- Zhang, Y., Feng, X., We, R. & Derynck, R. Receptor-associated Mad homologues synergize as effectors of the TGF-β response. *Nature* **383**, 168–172 (1996).
- Conlon, F. L. *et al.* A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* **120**, 1919–1928 (1994).
- Iannaccone, P. M., Zhou, X., Khokha, M., Boucher, D. & Kuehn, M. R. Insertional mutation of a gene involved in growth regulation of the early mouse embryo. *Dev. Dyn.* **194**, 198–208 (1992).
- Gu, Z. *et al.* The type I activin receptor ActRIB is required for egg cylinder organization and gastrulation in the mouse. *Genes Dev.* **12**, 844–857 (1998).
- Mountford, P. *et al.* Dicotronic targeting constructs: reporters and modifiers of mammalian gene expression. *Proc. Natl Acad. Sci. USA* **91**, 4303–4307 (1994).
- Hogan, B., Beddington, R., Costantini, F. & Lacy, E. In *Manipulating the Mouse Embryo* 51–73 (Cold Spring Harbor Laboratory Press, New York, 1994).
- Tam, P. L. P. & Behringer, R. R. Mouse gastrulation: the formation of a mammalian body plan. *Mech. Dev.* **68**, 3–25 (1997).
- Wilkinson, D. G., Bhatt, S. & Herrmann, B. G. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**, 657–659 (1990).
- Waldrip, W. R., Bikoff, E. K., Hoodless, P. A., Wrana, J. L. & Robertson, E. J. Smad2 signaling in extraembryonic tissues determines anterior-posterior polarity of the early mouse embryo. *Cell* **92**, 797–808 (1998).
- Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L. & Kuehn, M. R. Nodal is a novel TGF-β-like gene expressed in the mouse node during gastrulation. *Nature* **361**, 543–547 (1993).
- Varlet, I., Collignon, J. & Robertson, E. J. *Nodal* expression in the primitive endoderm is required for specification of the anterior axis during mouse gastrulation. *Development* **124**, 1033–1044 (1997).
- Collignon, J., Varlet, I. & Robertson, E. J. Relationship between asymmetric nodal expression and the direction of embryonic turning. *Nature* **381**, 155–158 (1996).
- Lowe, L. A. *et al.* Conserved left-right asymmetry of nodal expression and alterations in murine *situs inversus*. *Nature* **381**, 158–161 (1996).
- Oh, S. P. & Li, E. The signaling pathway mediated by the type IIB activin receptor controls axial patterning and lateral asymmetry in the mouse. *Genes Dev.* **11**, 1812–1826 (1997).
- Hoodless, P. A. *et al.* MADR1, a MAD-related protein that functions in BMP2 signaling pathways. *Cell* **85**, 489–500 (1996).
- Green, J. B., New, H. V. & Smith, J. C. Responses of embryonic *Xenopus* cells to activin and FGF are separated by multiple dose thresholds and correspond to distinct axes of the mesoderm. *Cell* **71**, 731–739 (1992).
- Jones, C. M., Kuehn, M. R., Hogan, B. L., Smith, J. C. & Wright, C. V. Nodal-related signals induce axial mesoderm and dorsalize mesoderm during gastrulation. *Development* **121**, 3651–3662 (1995).
- Lagna, G., Hata, A., Hemmati-Brivanlou, A. & Massagué, J. Partnership between DPC4 and SMAD proteins in TGF-β signalling pathways. *Nature* **383**, 832–836 (1996).
- Chen, X. *et al.* Smad4 and FAST-1 in the assembly of activin-responsive factor. *Nature* **389**, 85–89 (1997).
- Sirard, C. *et al.* The tumor suppressor gene *Smad4/Dpc4* is required for gastrulation and later for anterior development of the mouse embryo. *Genes Dev.* **12**, 107–119 (1998).
- Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
- Laird, P. W. *et al.* Simplified mammalian DNA isolation procedure. *Nucleic Acids Res.* **19**, 4293 (1991).
- Nakao, A. *et al.* Identification of Smad2, a human Mad-related protein in the transforming growth factor β signaling pathway. *J. Biol. Chem.* **272**, 2896–2900 (1997).
- Kaufman, M. H. *The Atlas of Mouse Development* (Academic, San Diego, 1992).
- Wilkinson, D. G. *In Situ Hybridization: A Practical Approach* (Oxford University Press, London, 1992).

Acknowledgements. We thank E. Robertson for *nodal^{lacZ}* mice; P. Ten Dijke for Smad2 antibody; A. Smith for pGTT1.8iresBgeo; P. Oh for help with whole-mount *in situ* hybridization; H. Lei, M. Okano and L. Yu for technical assistance; and Z. Gu and T. Gridley for comments on the manuscript. This work was funded by Bristol Myers-Squibb (E.L.) and a postdoctoral fellowship from the Japan Society for the Promotion of Science (M.N.).

Correspondence and requests for materials should be addressed to E.L. (e-mail: en@cvrc.mgh.harvard.edu).

Activation of human aortic smooth-muscle cells is inhibited by PPARα but not by PPARγ activators

Bart Staels*, Wolfgang Koenig†, Aïda Habib‡, Régine Merval§, Marilynne Lebret‡, Inés Pineda Torra*, Philippe Delerive*, Abdessamad Fadel*, Giulia Chinetti*, Jean-Charles Fruchart*, Jamila Najib*, Jacques Macclouf‡ & Alain Tedgui§

* U325 INSERM, Département d'Athérosclérose, Institut Pasteur, 1 Rue Calmette, 59019 Lille and the Faculté de Pharmacie, Université de Lille II, 59006 Lille, France

† Department of Internal Medicine II—Cardiology, University of Ulm, Robert Kochstrasse 8, 89081 Ulm, Germany

‡ U348 INSERM and § U141 INSERM and IFR 'Circulation Lariboisière', 41 Bd de la Chapelle, 75745 Paris Cédex 10, France

Peroxisome proliferator-activated receptors (PPARs) are key players in lipid and glucose metabolism and are implicated in metabolic disorders predisposing to atherosclerosis, such as dyslipidaemia and diabetes¹. Whereas PPARγ promotes lipid storage by regulating adipocyte differentiation, PPARα stimulates the β-oxidative degradation of fatty acids. PPARα-deficient mice show a prolonged response to inflammatory stimuli, suggesting that PPARα is also a modulator of inflammation². Hypolipidaemic fibrate drugs are PPARα ligands that inhibit the progressive formation of atherosclerotic lesions, which involves chronic inflammatory processes³, even in the absence of their atherogenic lipoprotein-lowering effect^{4,5}. Here we show that PPARα is expressed in human aortic smooth-muscle cells, which participate in plaque formation and post-angioplasty re-stenosis³. In these smooth-muscle cells, we find that PPARα ligands, and not PPARγ ligands, inhibit interleukin-1-induced production of interleukin-6 and prostaglandin and expression of cyclooxygenase-2. This inhibition of cyclooxygenase-2 induction occurs transcriptionally as a result of PPARα repression of NF-κB signalling. In hyperlipidaemic patients, fenofibrate treatment decreases the plasma concentrations of interleukin-6, fibrinogen and C-reactive protein. We conclude that activators of PPARα inhibit the inflammatory response of aortic smooth-muscle cells and decrease the concentration of plasma acute-phase proteins, indicating that PPARα in the vascular wall may influence the process of atherosclerosis and re-stenosis.

As inflammation involving activation of aortic smooth-muscle cells (SMC) is a hallmark of atherosclerosis³, we analysed the expression of PPARs in SMC. By using reverse transcription with the polymerase chain reaction (RT-PCR) and RNase protection analysis, we demonstrated the presence of PPARα messenger RNA in SMC (Fig. 1a, b). Although PPARγ mRNA was detectable by RT-PCR analysis (Fig. 1a), its expression was much lower than in adipose tissue and comparable to that in human liver, where its levels are low⁶. Western blot analysis using PPARα- and PPARγ-specific antibodies indicated that considerable amounts of PPARα protein were present, whereas only minute amounts of PPARγ protein were detected (Fig. 1c).

To determine whether PPARα interferes with the response of SMC to inflammatory cytokines, we analysed the influence of fibrates on interleukin (IL)-1-mediated activation of IL-6 production, a marker for SMC activation⁷. IL-1 alone increased SMC IL-6 secretion more than fivefold (Fig. 2a, b). Addition of PPARα activators (gemfibrozil, fenofibrate or Wy14643)^{2,8} by themselves did not influence production of IL-6 (Fig. 2a, and Supplementary

Information). However, fibrates prevented the IL-1-induced secretion of IL-6 in a dose-dependent manner (Fig. 2a, and Supplementary Information). This inhibition occurred at fibrate concentrations required for the induction of positive PPAR α -response genes⁹ and within the range of plasma concentrations found in humans^{10,11}. By contrast, the PPAR γ -specific ligand BRL49653 (ref. 12) did not affect basal or IL-1-induced IL-6 secretion (Fig. 2b), in agreement with the absence of appreciable PPAR γ expression in SMC (Fig. 1).

Next we determined whether PPAR α activators also interfere with prostaglandin (PG) production by SMC. Incubation of SMC with IL-1 increased 6-keto-PGF $_{1\alpha}$ secretion tenfold (Fig. 2c, d).

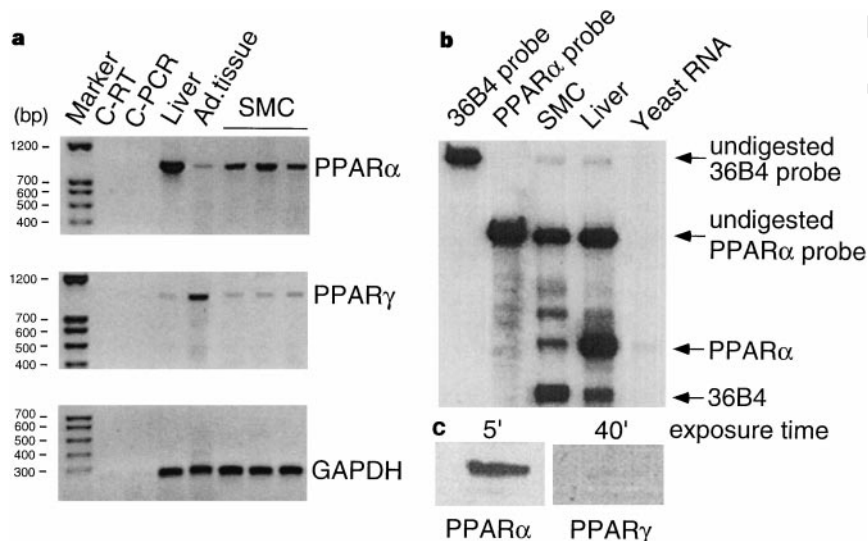


Figure 1 PPAR α , but not PPAR γ , is expressed in SMC. **a**, RT-PCR analysis of PPAR α , PPAR γ and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA. Human liver and adipose (ad.) tissue RNA served as controls for PPAR α and PPAR γ expression (relative molecular mass markers are shown in base pairs);

Fibrates, but not BRL49653, prevented the formation of 6-keto-PGF $_{1\alpha}$ by IL-1, whereas PPAR activators alone did not influence unstimulated 6-keto-PGF $_{1\alpha}$ secretion (Fig. 2c, d, and Supplementary Information).

Prostaglandin production from arachidonic acid is catalysed by the cyclooxygenase (COX) enzymes. Whereas COX-1 is constitutively expressed in most cells, COX-2 is rapidly induced upon stimulation by growth factors and cytokines. To determine whether PPAR α activators interfere with IL-1-induced PG production by inhibiting expression of cyclooxygenase, we did western blot analysis on extracts of SMC activated with IL-1 in the presence or absence of PPAR activators. We found that whereas COX-2 was

C-RT and C-PCR are negative controls for RT and PCR). **b**, RNase protection analysis of PPAR α mRNA. Human liver and yeast RNA served as positive and negative controls. **c**, Western blot analysis (50 μ g protein) of PPAR α and PPAR γ proteins.

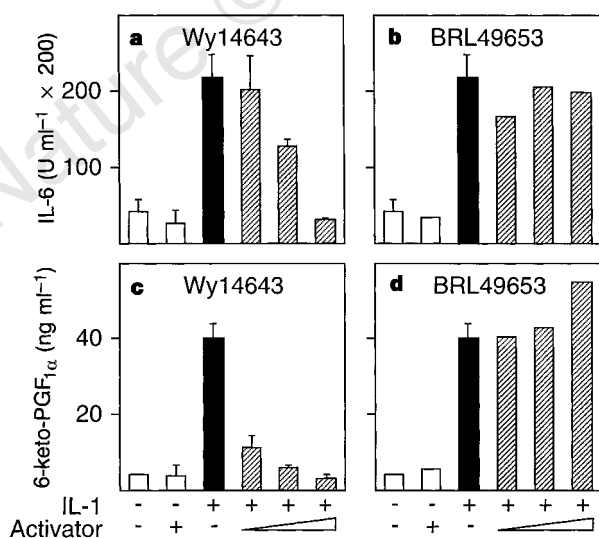


Figure 2 PPAR α , but not PPAR γ , activators inhibit the IL-1 mediated induction of IL-6 and 6-keto-PGF $_{1\alpha}$ secretion by SMC. **a**, **b**, IL-6 (ref. 27), and **c**, **d**, 6-keto-PGF $_{1\alpha}$ (ref. 28) secretion by SMC (passages 5–8) incubated for 24 h with recombinant human IL-1 β (1 ng ml $^{-1}$), Wy14643 (**a**, **c**; alone: 250 μ M; with IL-1: 50, 100 and 250 μ M) or BRL49653 (**b**, **d**; alone: 10 μ M; with IL-1: 0.1, 1 and 10 μ M). Results represent the mean $\pm$ s.e.m. of four (Wy14643) and two (BRL49653) experiments on SMC preparations from two different donors.

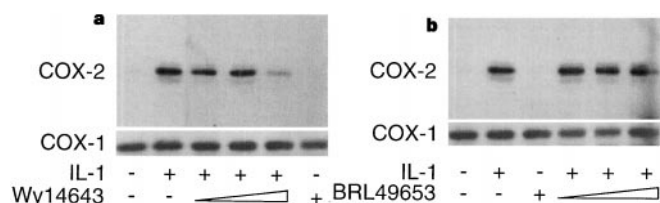


Figure 3 PPAR α activators, but not PPAR γ activators, inhibit IL-1-mediated induction of COX-2 protein in SMC. SMC were treated with IL-1 β (1 ng ml $^{-1}$), Wy14643 (**a**; alone: 250 μ M; with IL-1: 50, 100 and 250 μ M) or BRL49653 (**b**; alone: 10 μ M; with IL-1: 0.1, 1 and 10 μ M) and the concentration of COX-1 and COX-2 protein was measured by western blot analysis (20 μ g)²⁹.

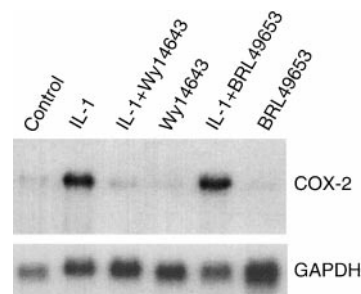


Figure 4 PPAR α activators, but not PPAR γ activators, inhibit the IL-1-mediated induction of COX-2 mRNA in SMC. SMC were incubated for 3 h with IL-1 β (1 ng ml $^{-1}$), Wy14643 (250 μ M) or BRL49653 (10 μ M), and COX-2 (ref. 30) and GAPDH mRNA were measured by northern blot analysis⁹.

barely detectable in unstimulated cells, activation by IL-1 markedly induced COX-2 (Fig. 3). Fibrates did not affect basal COX-2 protein levels, but prevented COX-2 induction by IL-1 (Fig. 3, and Supplementary Information). In contrast, BRL49653, at concentrations 10^3 -fold above its K_d for PPAR γ ¹², did not influence basal or IL-1-induced COX-2 protein levels (Fig. 3). COX-1 protein levels were not modulated by IL-1, fibrates or both together (Fig. 3). Like COX-2 protein, COX-2 mRNA was low in unactivated SMC but increased upon IL-1 stimulation (Fig. 4). Simultaneous addition of Wy14643, but not of BRL49653, prevented the induction of COX-2 mRNA by IL-1.

Inflammatory cytokines such as TNF- α , IFN- γ and IL-1 induce COX-2 gene transcription¹³. To determine whether PPAR α negatively interferes with the transcriptional activation of COX-2, the human COX-2 promoter¹⁴ was transiently transfected into Cos-1 cells in the presence or absence of a human PPAR α expression vector. Cells were subsequently treated with the phorbol ester PMA, an activator of COX-2 transcription, and/or Wy14643. COX-2 promoter activity was high (Fig. 5a), probably because of NF- κ B activation by serum components¹⁵. Wy14643 alone significantly decreased COX-2 promoter activity, probably as a result of activation of endogenous PPAR. PMA, an activator of the second messenger pathways used by inflammatory cytokines, significantly induced COX-2 promoter activity, which could be partially blocked by Wy14643. Co-transfection of human PPAR α prevented the PMA-mediated induction of COX-2 promoter activity, which was enhanced by its ligand Wy14643. These results indicate that PPAR α negatively regulates COX-2 gene transcription and prevents COX-2 induction by PMA.

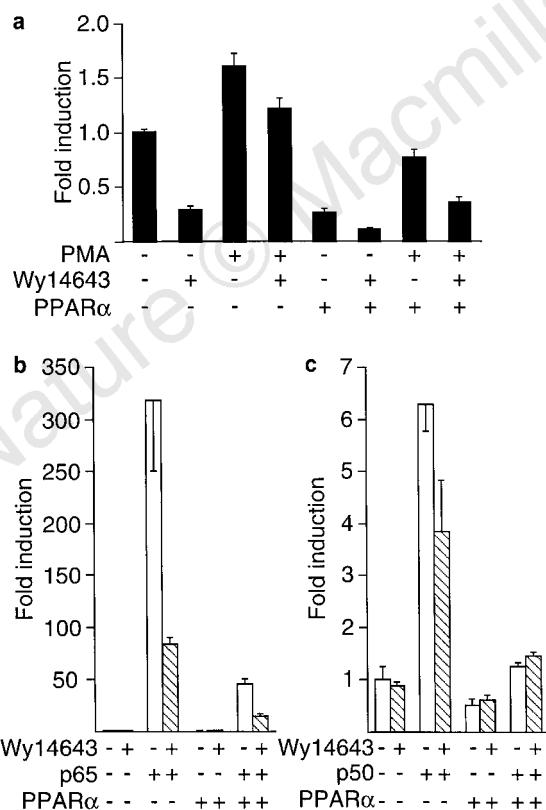


Figure 5 Fibrates inhibit the induction of human COX-2 promoter activity by PMA, p65/RelA and p50 via PPAR α . **a**, Cos-1 cells, transfected with the COX-2 gene promoter luciferase and PPAR α expression plasmids, were incubated (for 16 h) with Wy14643 (10 μ M) or vehicle (dimethylsulphoxide) and subsequently activated (6 h) with phorbol 12-myristate 13-acetate (PMA; 100 ng ml⁻¹) or Wy14643 (10 μ M). **b**, **c**, Cos-1 cells, transfected with the NF- κ B-response-element-driven luciferase reporter, PPAR α , p65/RelA and/or p50 expression plasmids, were incubated for 36 h with Wy14643 (10 μ M) or vehicle (dimethyl sulphoxide). Values represent the mean $\pm$ s.d.

The activation of COX-2 transcription by proinflammatory cytokines occurs by activation of the NF- κ B signalling pathway¹⁶. To determine whether PPAR α interferes with NF- κ B, we tested the influence of PPAR α on the transcriptional activation of an NF- κ B-response-element-driven reporter by the p65/RelA and p50 subunits of the NF- κ B transcription complex¹³. p65/RelA co-transfection resulted in a 300-fold activation of luciferase activity, which was significantly lowered by PPAR α and further enhanced by Wy14643 (Fig. 5b). Although p50-mediated activation was minor compared to p65/RelA, PPAR α co-transfection also prevented p50-mediated induction (Fig. 5c). Results were essentially identical in transfection experiments on the human COX-2 promoter (see Supplementary Information). Overexpression of p65/RelA resulted in a >25-fold induction of promoter activity, which was almost completely prevented by PPAR α and Wy14643.

To determine whether PPAR α activators influence the inflammatory response in humans, the influence of fenofibrate on plasma concentrations of acute-phase proteins was determined in patients with mild hyperlipidaemia with and without angiographically proven coronary artery disease (CAD). Plasma triglyceride (Fig. 6a), total and LDL (low-density lipoprotein) cholesterol (data not shown) concentrations were similar in both groups at baseline and decreased significantly upon fenofibrate treatment. However, plasma IL-6 concentrations were higher at baseline and decreased more markedly after fenofibrate treatment in patients with CAD compared with patients without CAD (Fig. 6c). Furthermore, fenofibrate significantly lowered plasma concentrations of fibrinogen and CRP, established risk factors for CAD^{17,18} whose production is controlled by cytokines (Fig. 6b, d). After discontinuation of fenofibrate therapy, IL-6 and triglyceride levels returned to baseline, whereas fibrinogen and CRP remained low in patients with CAD. These results indicate that fenofibrate improves the concentrations of acute-phase reactants in hyperlipidaemic patients and, most markedly, in those with CAD.

Therefore, PPAR α activators decrease the inflammatory response *in vitro* and *in vivo*. PPAR α activators exert their anti-inflammatory action, at least in part, by negatively regulating NF- κ B transcriptional activity. This anti-inflammatory activity may apply generally to PPAR family members as PPAR γ has been found to inhibit macrophage activation *in vitro*^{19,20}. Acute myocardial infarction is

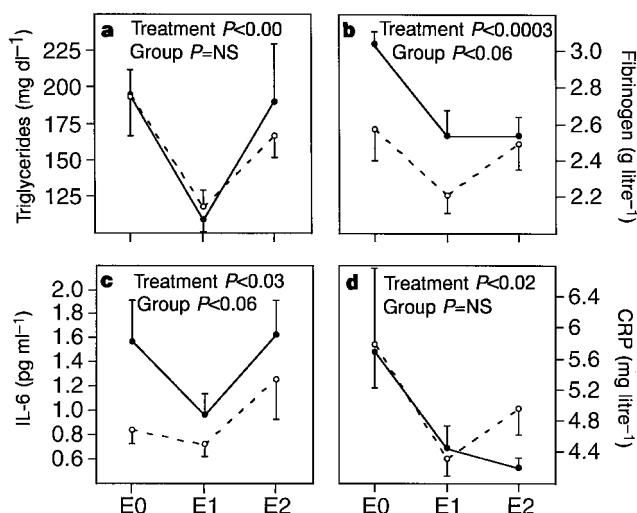


Figure 6 The PPAR α activator fenofibrate reduces the inflammatory response in patients with coronary atherosclerosis. Blood samples were taken from patients with documented coronary artery disease (filled circles) and controls (white circles) before (E0), after 4 weeks of fenofibrate (E1: 250 mg d⁻¹) and after 6 weeks of discontinuation of drug therapy (E2) and the following were measured: **a**, plasma triglyceride; **b**, fibrinogen; **c**, IL-6; and **d**, CRP. Values are expressed as mean $\pm$ s.e.m.

due to the rupture of unstable plaques followed by occlusive thrombosis. Unstable plaques are composed of a thin fibrous cap, a large lipid core and a high concentration of inflammatory cells and activated SMC³. In human SMC present in atherosclerotic lesions, the expression of cytokines, growth factors, cell-adhesion molecules, and proteins involved in extracellular matrix remodelling are all increased²¹. Furthermore, p65/RelA is activated in SMC, macrophages and endothelial cells in human atheroma^{22,23}, indicating that NF- κ B activation has a role in the dysregulation of vascular cell function. Therefore, in addition to their lipid-lowering properties, PPAR α activators have anti-inflammatory activity *in vivo*, so may have beneficial vascular effects in atherosclerosis and post-angioplasty re-stenosis, diseases in which SMC activation is important³. □

Methods

RNA analysis. RNase protection analysis was done using antisense human PPAR α (nucleotides (nt) 797–984) and control 36B4 (nt 202–371)²⁴ riboprobes. For RT-PCR amplification, we used the primers: 5'-GACGAATGC-CAAGATCTGAGAAAGC and 5'-CGTCTCCTTTGTAGTGCTGTCAGC (for PPAR α ; 948 bp); 5'-GGCAATTGAATGTCGTGTCTGTGGAGATAA and 5'-AGCTCCAGGGCTTGTAGCAGGTTGTCTTGA-3 (for PPAR γ ; 900 bp); 5'-ATGCAGCCCCGAATGCTCCTCATCGTGGCC and 5'-TTCTTGAGGC-CATGTGGGCCAT (for GAPDH; 239 bp).

Transient transfections. The human COX-2 gene promoter (nt -1,840/+123) or NF- κ B response element luciferase reporter plasmids (2 μ g) and human PPAR α , p65/RelA and/or p50 expression plasmids (1 μ g) were transfected⁹ and cells subsequently incubated in medium containing 0.2% fetal calf serum and the indicated activators. Luciferase activities were normalized to cellular protein and internal control plasmid pPGK β geobpA (500 ng) β -galactosidase activity²⁵.

Clinical study. Based on coronary angiography, mildly hyperlipidaemic patients were divided into 2 groups: one with significant CAD (at least one stenosis greater than 50% of the luminal diameter; $n = 19$) and one without CAD ($n = 19$). There were no significant differences in risk factors between the two groups. Fasting blood was drawn before (E0, where E is experimental time point), after 4 weeks of fenofibrate treatment (250 mg per day; E1) and 6 weeks after discontinuation (E2). Plasma triglycerides were measured enzymatically, fibrinogen by the Clauss method²⁶, CRP by nephelometry (Beckman, Germany) and IL-6 by enzyme-linked immunosorbent assay (R&D Systems).

Statistical analysis. Each parameter was measured as a univariate repeated-measures variable. The structure of the data consisted of a grouping factor (group) representing the patient groups and a repeated-measures factor (treatment) representing E0, E1 and E2. Effect estimates and statistical tests of group and treatment factors and their interactions were computed using the standard univariate repeated measures model.

Received 26 February; accepted 8 April 1998.

- Schoonjans, K., Martin, G., Staels, B. & Auwerx, J. Peroxisome proliferator-activated receptors, orphans with ligands and functions. *Curr. Opin. Lipidol.* **8**, 159–166 (1997).
- Devchand, P. R. *et al.* The PPAR α -leukotriene B₄ pathway to inflammation control. *Nature* **384**, 39–43 (1996).
- Ross, R. The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature* **362**, 801–809 (1993).
- Saitoh, K. *et al.* Anti-atheromatous effects of fenofibrate, a hypolipidemic drug. I: Anti-atheromatous effects are independent of its hypolipidemic effect in cholesterol-fed rabbits. *Folia Pharmacol. Japon.* **106**, 41–50 (1995).
- Ericsson, C. G. *et al.* Angiographic assessment of effects of bezafibrate on progression of coronary artery disease in young male postinfarction patients. *Lancet* **347**, 849–853 (1996).
- Auboeuf, D. *et al.* Tissue distribution and quantification of the expression of mRNAs of peroxisome proliferator-activated receptors and liver X receptor- α in humans. No alterations in adipose tissue of obese and NIDDM patients. *Diabetes* **46**, 1319–1327 (1997).
- Lopponow, H. & Libby, P. Proliferating or interleukin-1 activated human vascular smooth muscle cells secrete copious interleukin 6. *J. Clin. Invest.* **85**, 731–738 (1990).
- Forman, B. M., Chen, J. & Evans, R. M. Hypolipidemic drugs, polyunsaturated fatty acids, and eicosanoids are ligands for peroxisome proliferator-activated receptors α and δ . *Proc. Natl Acad. Sci. USA* **94**, 4312–4317 (1997).
- Vu-Dac, N. *et al.* Fibrates increase human apolipoprotein A-II expression through activation of the peroxisome proliferator-activated receptor. *J. Clin. Invest.* **96**, 741–750 (1995).
- Balfour, J. A., McTavish, D. & Heel, R. C. Fenofibrate, a review of its pharmacodynamic and pharmacokinetic properties and therapeutic use in dyslipidaemia. *Drugs* **40**, 260–290 (1990).
- Todd, P. A. & Ward, A. Gemfibrozil, a review of its pharmacodynamic and pharmacokinetic properties, and therapeutic use in dyslipidaemia. *Drugs* **36**, 314–339 (1988).
- Lehmann, J. M. *et al.* An antidiabetic thiazolidinedione is a high affinity ligand for Peroxisome Proliferator-Activated Receptor γ (PPAR γ). *J. Biol. Chem.* **270**, 12953–12956 (1995).
- Newton, R., Kuitert, L. M. E., Bergmann, M., Adcock, I. M. & Barnes, P. J. Evidence for involvement of NF- κ B in the transcriptional control of COX-2 gene expression by IL-1 β . *Biochem. Biophys. Res. Commun.* **237**, 28–32 (1997).

- Kutcher, W. *et al.* Prostaglandin H synthase 2 is expressed abnormally in human colon cancer: evidence for a transcriptional effect. *Proc. Natl Acad. Sci. USA* **93**, 4816–4820 (1996).
- Xie, W. & Herschman, H. R. Transcriptional regulation of prostaglandin synthase 2 gene expression by platelet-derived growth factor and serum. *J. Biol. Chem.* **271**, 31742–31748 (1996).
- Yamamoto, K., Arakawa, T., Ueda, N. & Yamamoto, S. Transcriptional roles of nuclear factor κ B and nuclear factor-interleukin-6 in the tumor necrosis factor α -dependent induction of cyclooxygenase-2 in MC3T3-E1 cells. *J. Biol. Chem.* **270**, 31315–31320 (1995).
- Liuzzo, G. *et al.* The prognostic value of C-reactive protein and serum amyloid A protein in severe unstable angina. *N. Engl. J. Med.* **331**, 417–424 (1994).
- Thompson, S. G., Kienast, J., Pyke, S. D. M., Haverkate, E. & van de Loo, J. C. W. Hemostatic factors and the risk of myocardial infarction or sudden death in patients with angina pectoris. *N. Engl. J. Med.* **332**, 635–641 (1995).
- Ricote, M., Li, A. C., Willson, T. M., Kelly, C. J. & Glass, C. K. The peroxisome proliferator-activated receptor- γ is a negative regulator of macrophage activation. *Nature* **391**, 79–82 (1998).
- Jiang, C., Ting, A. T. & Seed, B. PPAR- γ agonists inhibit production of monocyte inflammatory cytokines. *Nature* **391**, 82–86 (1998).
- Galis, Z. S., Sukhova, G. K., Lark, M. W. & Libby, P. Increased expression of matrix metalloproteinases and matrix degrading activity in vulnerable regions of human atherosclerotic plaques. *J. Clin. Invest.* **94**, 2493–2503 (1994).
- Brand, K. *et al.* Activated transcription factor nuclear factor-kappa B is present in the atherosclerotic lesion. *J. Clin. Invest.* **97**, 1715–1722 (1996).
- Bourcier, T., Sukhova, G. & Libby, P. The nuclear factor κ -B signalling pathway participates in dysregulation of vascular smooth muscle cells *in vitro* and in human atherosclerosis. *J. Biol. Chem.* **272**, 15817–15824 (1997).
- Rich, B. E. & Steitz, J. A. Human acidic ribosomal phosphoprotein P0, P1 and P2: analysis of cDNA clones, *in vitro* synthesis and assembly. *Mol. Cell. Biol.* **7**, 4065–4074 (1987).
- Soriano, P., Montgomery, C., Geske, R. & Bradley, A. Targeted disruption of the *c-src* proto-oncogene leads to osteopetrosis in mice. *Cell* **64**, 693–702 (1991).
- Clauss, A. Gerinnungsphysiologische Schnellmethode zur Bestimmung des Fibrinogens. *Acta Haematol.* **17**, 237–246 (1957).
- Belmin, J. *et al.* Increased production of tumor necrosis factor and interleukin 6 by the arterial wall of aged rats. *Am. J. Physiol.* **268**, H2288–H2293 (1995).
- Pradell, F., Gras, J. & Maclouf, J. Enzyme immunoassay of eicosanoids using acetylcholine esterase as a label: an alternative to radioimmunoassay. *Analyt. Chem.* **57**, 1170–1173 (1985).
- Habib, A. *et al.* Demonstration of an inducible cyclooxygenase in human endothelial cells using antibodies raised against the carboxyl-terminal region of the cyclooxygenase-2. *J. Biol. Chem.* **268**, 23448–23454 (1993).
- Jones, D. A., Carlton, D. P., McIntyre, M. P., Zimmerman, G. A. & Prescott, S. M. Molecular cloning of human prostaglandin endoperoxide synthase type II and demonstration of expression in response to cytokines. *J. Biol. Chem.* **268**, 9049–9054 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We acknowledge the technical contribution of O. Vidal, P. Poulin and B. Derudas, and thank C. Haegeman for p50 and p65/RelA, S. Prescott for human COX-2 promoter plasmids, and M. Sund for help with statistical analysis. This work was supported by grants from the Association pour la Recherche sur le Cancer (to J.M.), CNAMTS/INSERM (to A.T.) and the Région Nord-Pas de Calais. B.S. and J.M. are members of the CNRS.

Correspondence and requests for materials should be addressed to B.S. (e-mail: Bart.Staels@pasteur-lille.fr).

Modulation of AMPA receptor unitary conductance by synaptic activity

Tim A. Benke^{*†‡}, Andreas Lüthi^{*†‡}, John T. R. Isaac^{*} & Graham L. Collingridge^{*}

^{*} Department of Anatomy, University of Bristol, University Walk, Bristol BS8 1TD, UK

[†] These authors contributed equally to this work

Activity-dependent alteration in synaptic strength is a fundamental property of the vertebrate central nervous system and is thought to underlie learning and memory. The most extensively studied model of activity-dependent synaptic plasticity is long-term potentiation (LTP) of glutamate-responsive (glutamatergic) synapses, a widespread phenomenon involving multiple mechanisms¹. The best characterized form of LTP occurs in the CA1 region of the hippocampus, in which LTP is initiated by transient activation of NMDA (N-methyl-D-aspartate) receptors and is expressed as a persistent increase in synaptic transmission through AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate) receptors². This increase is due, at least in part, to a postsynaptic modification of AMPA-receptor function³; this modifi-

[†] Present addresses: Department of Pediatrics, Baylor College of Medicine, Houston, Texas 77030, USA (T.A.B.); Brain Research Institute, University of Zurich, CH-8029, Zurich, Switzerland (A.L.).

cation could be caused by an increase in the number of receptors, their open probability, their kinetics or their single-channel conductance. Here we show that the induction of LTP in the CA1 region of the hippocampus is often associated with an increase in single-channel conductance of AMPA receptors. This shows that elementary channel properties can be rapidly modified by synaptic activity and provides an insight into one molecular mechanism by which glutamatergic synapses can alter their strength.

We tested the feasibility of applying peak-scaled, non-stationary fluctuation analysis (non-SFA) to estimate the single-channel conductance (γ) of AMPA receptors^{4–6} at synapses on CA1 pyramidal

neurons by using a multicompartmental model incorporating a range of electrotonic properties^{7–9}, a kinetic scheme for AMPA-receptor channels with randomized mean open time (parameter range, 0.5–4 ms)^{7,10}, and a simulated patch electrode. Increases in either γ (8–40 pS) or the number of synaptically activated channels (20–160 channels) could be differentiated using non-SFA. The estimate of γ was insensitive to changes in peak open probability (0.3–0.8), the total channel number (up to a fourfold increase), the number of synaptic sites activated with random latency, glutamate concentration (0.5–10 mM) and duration (2 μ s to 1 ms)¹¹ and spine morphology (neck diameter, 0.55–1 μ m; ratio of head to neck diameters, 1–5.5). Increasing the channel mean open time (0.5–4 ms) changed the estimate of γ ; however, this change in γ was always accompanied by marked alterations in the decay time of the excitatory postsynaptic currents (EPSCs). Estimates of γ were affected by changes in membrane capacitance (C_m , 0.65–1.0 μ F cm⁻²), membrane resistivity (R_m , 38–150 k Ω cm²), axial resistivity (R_i , 100–294 Ω cm) and access resistance (R_A , 0.5–60 M Ω). However, alterations of >12% in γ were always associated with marked changes in the response of the model to input step voltages. Estimates of γ were also affected by the distance between the electrode and the synaptic site (0.5–40 μ m) because of dendritic filtering. Within all these parameter ranges, however, non-SFA provided an accurate measure of changes in γ (that is, for a given

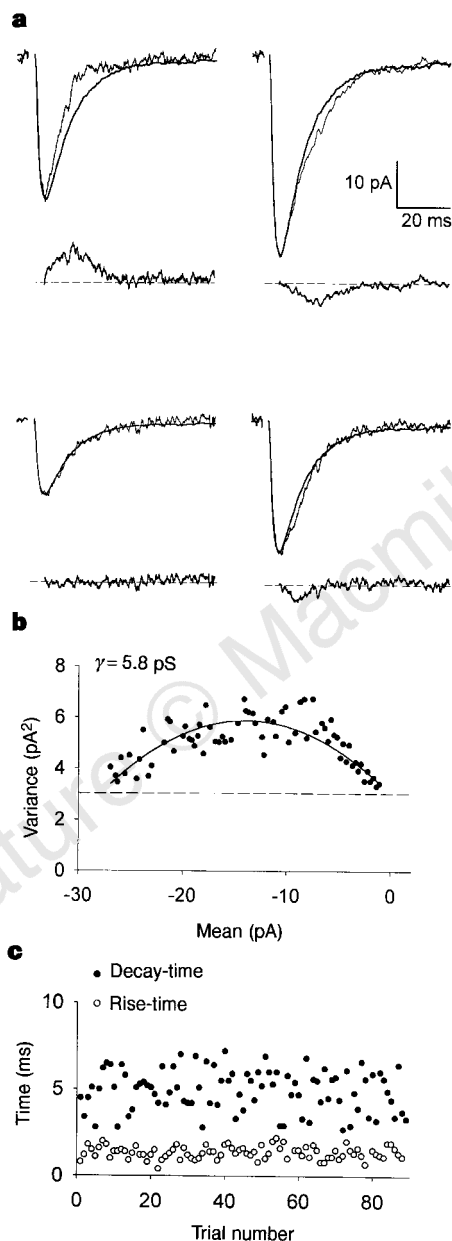


Figure 1 Non-stationary fluctuation analysis (non-SFA) of EPSCs recorded from CA1 dendrites. **a**, Top traces in each pair, four EPSCs (thin lines) and the superimposed scaled mean of 89 EPSCs (thicker lines). Bottom traces, the difference between each EPSC and the scaled mean (dashed lines represent no difference). **b**, Corresponding current-variance relationship; variance of the fluctuation of the decays of the individual EPSCs about the mean is plotted as a function of the mean EPSC decay amplitude. Solid line is the parabolic fit from which γ is estimated. Dashed line represents the background variance. **c**, Rise times (20–80%, open circles) and decay times (62%, filled circles) for all individual responses in this time period.

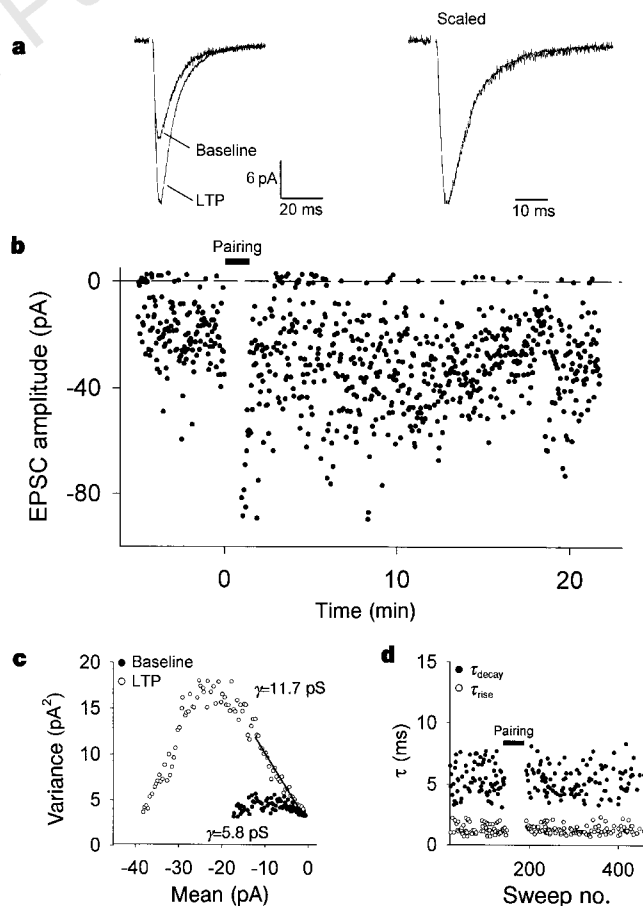


Figure 2 LTP is associated with an increase in AMPA-receptor conductance (γ). **a**, Mean EPSCs during baseline and LTP superimposed (left) and scaled (right; expanded timescale). **b**, All EPSC amplitudes during this experiment. **c**, Current-variance relationships from baseline (filled circles) and during LTP (open circles) for this cell. Solid lines represent the data fit from which γ was estimated (for this and all following figures). **d**, Rise times (20–80%, open circles) and decay times (62%, filled circles) for all EPSCs used for non-SFA. There was no significant change in the variance of rise times and decay times following the induction of LTP ($n = 21$).

set of parameters, if input γ was increased then the value of γ derived by non-SFA was also increased by the same proportion).

The modelling indicates that non-SFA can be used to accurately detect changes in γ in synapses remote from the soma. However, to minimize attenuation of γ , and hence maximize the signal-to-noise ratio, it was necessary to make stable recordings with low R_A as close as possible to the synapse. Therefore, we obtained visually guided, whole-cell voltage-clamp recordings from proximal apical dendrites of CA1 neurons¹² and activated synapses near to the recording site by minimal stimulation. In 53 recordings in which relatively low access resistances were obtained ($R_A = 34 \pm 1 \text{ M}\Omega$, range = 19–56 $\text{M}\Omega$), EPSCs exhibited rapid kinetics ($\tau_{\text{rise}} = 1.41 \pm 0.06 \text{ ms}$, $\tau_{\text{decay}} = 7.2 \pm 0.3 \text{ ms}$; $n = 53$; see Methods) compared with somatically recorded currents from CA1 neurons^{13–15} and non-SFA could be performed (Fig. 1). We obtained an estimated mean value for γ at CA1 pyramidal cell synapses of $7.7 \pm 0.7 \text{ pS}$, with a range of 1.5–22.3 pS ($n = 53$). This mean, and range, of γ values is similar to that reported for cerebellar granule cell synapses ($11 \pm 2 \text{ pS}$)⁶ and for patches excised from CA1 pyramidal cell dendrites ($10.2 \pm 0.8 \text{ pS}$)⁸. There was no correlation between γ and average EPSC amplitude, potency (defined as the amplitude of EPSCs excluding failures)¹⁴, τ_{rise} , τ_{decay} , R_A or background variance (data not shown). This indicates that the range of values for γ seen here is mainly due to variations in AMPA-channel conductance amongst synapses.

Pairing-induced LTP resulted in stable potentiation of EPSC amplitude ($209 \pm 19\%$; $n = 21$) that persisted for as long as dendritic recordings were maintained (up to 30 min after pairing). There was no change in either τ_{rise} or τ_{decay} after pairing (see scaled EPSCs in Fig. 2; baseline: $\tau_{\text{rise}} = 1.6 \pm 0.1 \text{ ms}$, $\tau_{\text{decay}} = 7.3 \pm 0.4 \text{ ms}$, $R_A = 35 \pm 1 \text{ M}\Omega$; LTP: $\tau_{\text{rise}} = 1.6 \pm 0.1 \text{ ms}$, $\tau_{\text{decay}} = 7.7 \pm 0.4 \text{ ms}$; $R_A = 36 \pm 1 \text{ M}\Omega$; $n = 21$). Non-SFA was performed to estimate γ before and after LTP induction (Fig. 2), using a minimum of 15 EPSCs (mean number: 37 ± 3 ; range: 15–104). On average there was an increase in γ during LTP of $184 \pm 20\%$ ($P < 0.001$; $n = 21$). However, these cells fell into two distinct populations (Fig. 3). In 13 cells, LTP was associated with a robust increase in γ ($236 \pm 19\%$; $P < 0.001$; $n = 13$). Analysis of sequential 10-min periods following LTP showed that this increase was stable for the duration of the recording. In most neurons, the increase in γ was sufficient to account for LTP (Fig. 4c). In the other eight cells there

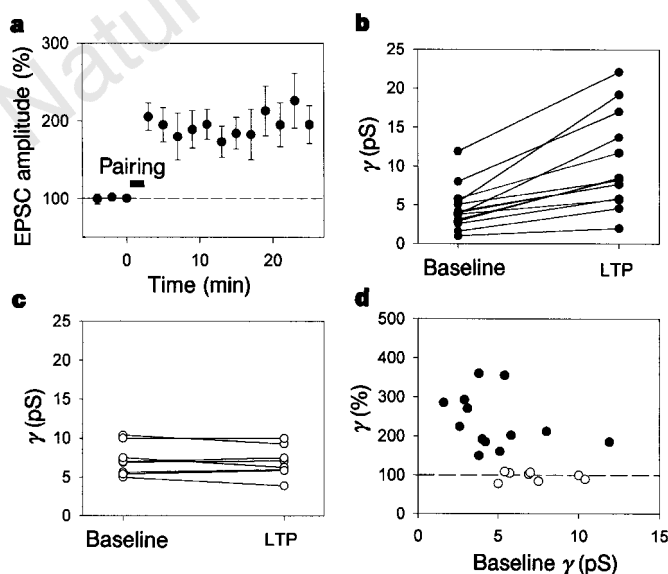


Figure 3 A summary of changes in γ associated with LTP. **a**, EPSC amplitude (2-min bins) plotted against time for all 21 LTP experiments. **b**, **c**, Absolute values of γ for all cells in which **(b)** γ increased or **(c)** γ did not change. **d**, Relationship between γ increase and baseline γ for all neurons. Open circles, no change in γ ; filled circles, increase in γ .

was no change in γ ($97 \pm 4\%$). There was a significant difference between those neurons showing an increase in γ and those that did not with respect to baseline γ values ($4.8 \pm 0.8 \text{ pS}$, $n = 13$; versus $7.2 \pm 0.7 \text{ pS}$, $n = 8$, respectively; $P < 0.05$; Fig. 3d), but there was no significant difference with respect to baseline potency (that is, potency of EPSCs before LTP), EPSC kinetics, R_A , input resistance and background variance (not shown).

We performed interleaved control experiments in which EPSC amplitude was changed using methods that theoretically should not alter γ (Fig. 4a–c). Paired-pulse facilitation (interpulse interval, 50 ms) did not affect γ even though there was a comparable amount of potentiation to that seen in the LTP experiments (first response $\gamma = 8.0 \pm 1.0 \text{ pS}$; second response $\gamma = 7.5 \pm 0.9 \text{ pS}$, per cent amplitude = 185 ± 17 ; per cent potency = 138 ± 5 , $n = 8$; Fig. 4a, c). Hyperpolarization increased the single channel current, because of the increased driving force, but γ did not change (depolarized potential: $\gamma = 8.2 \pm 3.4 \text{ pS}$; hyperpolarized potential: $\gamma = 7.7 \pm 3.4 \text{ pS}$; per cent amplitude = 140 ± 9 , per cent potency = 146 ± 15 , $n = 4$; Fig. 4b, c). Reduction of EPSC amplitude with NBQX (2,3-dihydroxy-6-nitro-7-sulphamoylbenzo quinoxaline) ($0.2 \mu\text{M}$) had no effect on γ (pre-NBQX: $\gamma = 4.6 \pm 1.9 \text{ pS}$; NBQX: $\gamma = 4.9 \pm 1.8 \text{ pS}$; per cent amplitude = 55 ± 5 ; $n = 3$; Fig. 4c). Addition of D-AP5 (D-2-amino-5-phosphonopentanoate) ($50 \mu\text{M}$) during baseline or after induction of LTP had no effect on γ (baseline: pre-D-AP5, $\gamma = 6.8 \pm 0.2 \text{ pS}$; D-AP5, $\gamma = 7.1 \pm 0.3 \text{ pS}$;

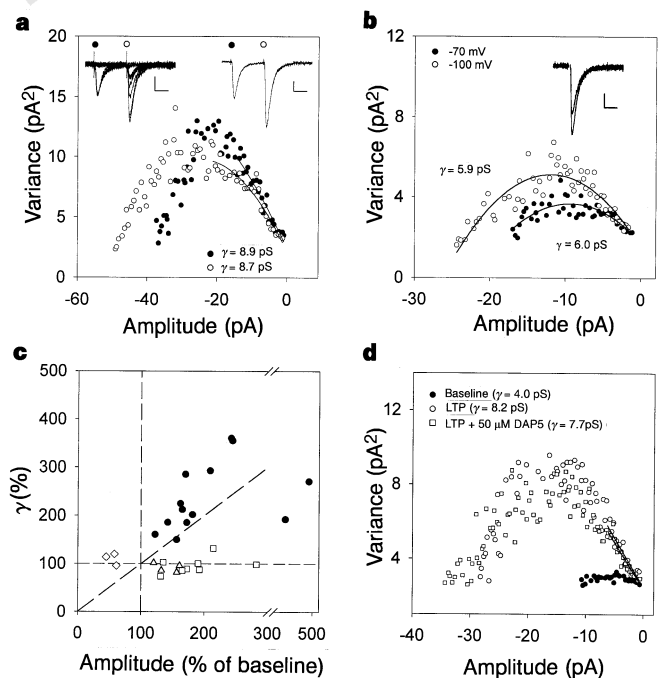


Figure 4 Increase in γ is specific to LTP. **a**, **b**, Current-variance relationship for **(a)** a paired-pulse facilitation experiment (inset left, eight consecutive superimposed traces; inset right, epoch average) and **(b)** a holding potential experiment (inset, mean EPSCs at -100 mV and -70 mV). Inset scale bars represent 20 pA , 5 pA and 5 pA , respectively (y -axis) versus 20 ms . **c**, Change in γ plotted as a function of change in mean EPSC amplitude for all paired-pulse experiments (squares), for all holding potential experiments (triangles), for all experiments with NBQX (diamonds), and for all LTP experiments that exhibited a conductance increase (filled circles). Diagonal dashed line indicates a 1:1 relationship. The greater change in γ than in amplitude during LTP seen in most cells may be due to dendritic filtering attenuating the peak response to a greater degree than the tail (over which γ is estimated); this was also observed in the model. However, when all cells were considered together this effect was not significant ($18 \pm 30\%$, $n = 13$). **d**, Current-variance plot showing the lack of effect of subsequent application of D-AP5 ($50 \mu\text{M}$) on the increase in γ caused by LTP.

$n = 3$; not illustrated; LTP: pre-D-AP5, $\gamma = 5.8 \pm 0.6$ pS; LTP + D-AP5, $\gamma = 6.8 \pm 0.7$ pS; $n = 6$). Importantly, in four of these experiments, induction of LTP resulted in an increase in γ , and this was not significantly affected by subsequent addition of D-AP5 (baseline: 3.5 ± 0.3 pS; LTP pre-D-AP5: 6.3 ± 0.7 pS; LTP + D-AP5: 7.6 ± 0.6 pS; baseline versus LTP: $P < 0.05$; $n = 4$; Fig. 4d). Finally, in experiments in which pairing failed to induce LTP, presumably because of washout¹³ ($n = 4$), or in which LTP was blocked by the application of D-AP5 (50 μ M) during induction ($n = 2$), γ did not change (baseline $\gamma = 6.0 \pm 0.7$ pS, post-pairing $\gamma = 5.6 \pm 0.9$ pS; $n = 6$; not illustrated).

Failures analysis of the neurons exhibiting LTP, using visual identification of failures (Fig. 5a), showed that potentiation was associated with large changes in potency (from 12.3 ± 1.4 pA to 23.6 ± 2.5 pA, $193 \pm 14\%$; $n = 21$; $P < 0.001$; Fig. 5b) and small changes in success rate (from 0.67 ± 0.05 to 0.73 ± 0.05 ; $n = 21$; $P < 0.01$; Fig. 5c). As shown in Fig. 5d, in most neurons that exhibited an increase in γ (8 of 13) success rate changed by $< 10\%$.

Our main finding is that synaptic activity can result in an increase in the single-channel conductance of AMPA receptors at CA1 synapses. This increase in γ is specifically related to the induction of LTP, as it was not observed during manipulations that changed the synaptic response in other ways (paired-pulse facilitation, increasing driving force or addition of NBQX). Several possible mechanisms could explain the increase in γ . First, release probability could be increased at previously presynaptically silent synapses with higher-conductance receptors; this would increase the average value of γ . However, as we found that γ increased in the absence of an appreciable change in the success rate of most cells, this explanation is unlikely^{15,16}. Second, higher-conductance pre-assembled AMPA receptors that would confer a greater average γ might be inserted; for example, such AMPA receptors might contain unedited GluR2 or lack GluR2 receptors altogether¹⁷. However, as the increase in γ was sufficient, at least for most cells, to account for LTP, it is unlikely that there was a net increase in the number of AMPA receptors at the synapse. Therefore, to keep the number of

receptors constant, the increase in high-conductance AMPA receptors would need to be counterbalanced by an equivalent decrease in low-conductance AMPA receptors or by a compensatory decrease in the channel open probability (P_{open}). Third, receptors that are already inserted into the synapse might be modified. It is known that AMPA-receptor channels can be rapidly modulated, for example by protein kinase A activators^{18,19}, and this is associated with a change in channel kinetics¹⁸. However, a change in mean open time, deactivation or desensitization cannot account for our findings as there was no change in EPSC decay kinetics. AMPA-receptor-gated channels can exist in multiple conductance states^{17,20,21} and the estimate for synaptic channel conductance will be a weighted mean of any different levels. Therefore, the increase in synaptic γ could be due to an increase in the relative time that AMPA receptors spend in the high-conductance states, perhaps as a result of a direct phosphorylation²².

Although an increase in conductance accounts for LTP in most cells, other processes must also be able to contribute, as in some cells γ did not change and in other neurons there were also changes in success rate. These effects could also be presynaptic or postsynaptic in origin²³. For example, the rapid insertion of AMPA receptors with the same average γ to those already inserted into the synaptic membrane could account for the cells in which no change in γ was observed²⁴ and, if this also happened at silent synapses^{25,26}, could also account for the small changes in success rates that were sometimes observed. Another possible mechanism is an increase in P_{open} of existing AMPA-receptor channels²⁷.

All of our results could be explained by Ca^{2+} permeation through NMDA channels, resulting in activation of kinases¹ and leading to incorporation of new receptors and/or modification of the conductance properties of existing ones. The correlation between the increase in γ and the baseline γ indicates that the initial state of the AMPA receptors may determine how synaptic efficiency is altered. The finding that synaptic activity can modify the single channel conductance properties of the AMPA-receptor complement provides a fundamental mechanism for the alteration of synaptic efficiency. □

Methods

Modelling. The passive electrical properties of a neuron were simulated with a mathematical model using the formulation of Hines²⁸. This model was 25–29 compartments with, in most simulations, 7 compartments devoted to a single dendritic spine²⁹. A correction factor of 2.25 was applied to account for the membrane capacitance of spines not explicitly modelled⁹. AMPA-receptor channels were modelled with a kinetic scheme⁷. The peak open probability of the scheme was altered by varying the opening rate constant, α , from 3.0×10^3 to $4.24 \times 10^5 \text{ s}^{-1}$. Modelled activation of these receptors resulted in membrane current detected by the simulated electrode when this scheme was added to the compartmental model. Receptors were placed on either the end of an explicitly modelled dendritic spine or at multiple locations along the dendrite. When simulating multiple synaptic sites, individual sites were randomly activated with random latencies (0–1 ms).

Electrophysiology. Transverse hippocampal slices (400 μ m) obtained from rats aged 13–15 days were perfused with a solution containing (in mM): 124 NaCl, 3 KCl, 1.25 NaH_2PO_4 , 26 NaHCO_3 , 2 CaCl_2 , 1 MgSO_4 , 15 glucose, 2 ascorbic acid, 0.05 picrotoxin, saturated with 95% O_2 /5% CO_2 at room temperature (23–25 °C). Dendrites were visualized using infrared illumination and differential interference contrast (DIC) optics and approached under visual control. Whole-cell dendritic recordings were obtained at -70 mV using patch electrodes (6–10 M Ω) filled with a solution containing (in mM): 135 CsMeSO₄, 8 NaCl, 10 HEPES, 0.5 EGTA, 4 Mg-ATP, 0.3 Na-GTP, 5 QX-314, pH 7.25, 285 mOsm. Schaffer collateral–commissural fibres were stimulated at 0.3–1 Hz using an insulated platinum monopolar electrode (5 M Ω resistance, WPI), positioned 20–40 μ m from the dendrite. To obtain sufficient EPSCs to be able to estimate γ during the baseline before washout of LTP¹³, stimulus intensity was set so that success rate (1 – failure rate) was relatively high (baseline success rate = 0.67 ± 0.05 , range 0.25–0.97, $n = 21$). LTP was induced by delivering 30–60 stimulations at a holding potential of -10 mV

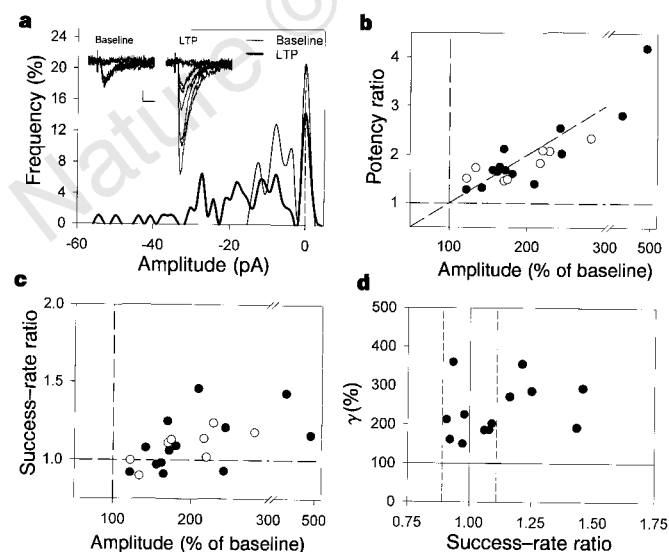


Figure 5 Failures analysis of LTP of dendritically recorded EPSCs. **a**, Amplitude histogram from a single experiment (bin width, 1 pA). Inset, eight superimposed consecutive traces for baseline (left panel) and during LTP (right panel; scale bar: 5 pA (y-axis)/10 ms). **b**, **c**, The ratio (LTP: baseline) for **(b)** potency and **(c)** success rate (1 – failure rate) as a function of EPSC amplitude (per cent of baseline) for cells that exhibited an increase (filled circles) or no change (open circles) in γ during LTP. **d**, Change in γ as a function of change in success rate for all neurons that exhibited an increase in γ with LTP. The dashed lines indicate a change in success rate of $\pm 10\%$.

to +10 mV (pairing) without interrupting stimulation or changing stimulation frequency. Recordings were made using an Axopatch 1D; signals were filtered at 5 kHz (8 pole Bessel filter), digitized at 10 kHz and stored on computer. EPSC amplitude, input resistance and whole-cell conductance were analysed and displayed online. Access resistance (R_A) was estimated by measuring the peak amplitude of the fast whole-cell capacitance current in response to a -1 mV step applied to the cell during each sweep. The amplitude was estimated by fitting the capacitance transient with a double exponential (from 0.5 ms after the peak) and determining the current at the beginning of the step. There was no change in the membrane response to a voltage step following the induction of LTP.

Analysis. Responses were carefully selected for analysis by visual inspection based on the following criteria: first, fast rise time allowing for precise alignment; second, no contamination of the EPSC or preceding baseline by sudden fluctuation in holding current or the presence of spontaneous activity; third, a latency of no more than 3 ms after stimulation. Data were analysed using peak-scaled non-SFA⁴⁻⁶, as follows (Fig. 1): an ensemble of simulated currents (125–500) or EPSCs (15–104) was aligned by the point of maximal rise, and averaged. The average response was scaled to the peak and subtracted from individual responses. The variance of the fluctuation around the mean was calculated for 10–100 bins, of equal current decrement, from the peak of the response until 5–6 times τ_{decay} (see below). The binned variance was plotted against the mean current amplitude and the single-channel current was estimated by fitting the data using a least-squares algorithm according to $\sigma^2 = iI - I^2/N + b_1$, where σ^2 is the variance, I is the mean current, N is the number of channels activated at the peak of the mean current, i is the single channel current and b_1 is the background variance³⁰. The single-channel conductance, γ , is then $\gamma = i/V$, where V is the driving force (holding potential minus assumed reversal potential of 0 mV).

To obtain the most accurate estimate for γ , different portions of the data were fitted (from 0 to either 20, 30, 50 or 100% of maximum amplitude). The goodness of fit was assessed using a least-squares algorithm and the fit of the rising phase of the current–variance plot at the low-current-amplitude values was also assessed by eye. Within a given experimental manipulation, the same fraction of the data was always used for the fit. All the estimates in which the initial rising phase of the current–variance plot were well fitted produced very similar values of γ , as has been reported⁴. Finally, a full blind re-analysis was also done by a second experimenter in most of the LTP experiments. This yielded similar values of γ in all cells.

To determine the minimum number of EPSCs required for a reliable estimate of γ , we compared estimates of γ from increasing numbers of EPSCs (5, 10, 15, 20, 25, 40 and 50) to the estimate of γ from the entire data set, for five randomly selected neurons; an accurate estimate of γ was obtained using data sets of 15 EPSCs or more. EPSC amplitude was estimated by measuring the difference between the average current over two time windows of equal length, one immediately before the stimulus artefact and the other centred on the peak of the response.

We obtained estimates of the rise and decay times of individual currents by measuring the 20–80% rise time and 62% decay time, respectively (referred to as 'rise time' and 'decay time'). For mean currents, we calculated EPSC kinetics using the time constants of single exponential fits (referred to as ' τ_{rise} ' and ' τ_{decay} '). For display of individual EPSC traces in the figures, the stimulus artefact was digitally subtracted using an average of identified failures. Data are expressed as percentage of baseline (that is, 100% indicates no change). All values are expressed as mean $\pm$ s.e.m. Statistical significance was assessed using two-tailed paired or unpaired Student's t -tests as appropriate ($P < 0.05$ considered significant).

Received 23 December 1997; accepted 9 April 1998.

1. Bliss, T. V. P. & Collingridge, G. L. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
2. Collingridge, G. L., Kehl, S. J. & McLennan, H. Excitatory amino acids in synaptic transmission in the Schaffer collateral-commissural pathway of the rat hippocampus. *J. Physiol. (Lond.)* **334**, 33–46 (1983).
3. Davies, S. N., Lester, R. A. J., Reymann, K. G. & Collingridge, G. L. Temporally distinct pre- and post-synaptic mechanisms maintain long-term potentiation. *Nature* **338**, 500–503 (1989).
4. Traynelis, S. F., Silver, R. A. & Cull-Candy, S. G. Estimated conductance of glutamate receptor channels activated during EPSCs at the cerebellar mossy fiber-granule cell synapse. *Neuron* **11**, 279–289 (1993).
5. De Koninck, Y. & Mody, I. Noise analysis of miniature IPSCs in adult rat brain slices: Properties and modulation of synaptic GABA_A receptor channels. *Neurophysiol.* **71**, 1318–1335 (1994).
6. Silver, R. A., Cull-Candy, S. G. & Takahashi, T. Non-NMDA glutamate receptor occupancy and open

- probability at a rat cerebellar synapse with single and multiple release sites. *J. Physiol. (Lond.)* **494**, 231–250 (1996).
7. Jonas, P., Major, G. & Sakmann, B. Quantal components of unitary EPSCs at the mossy fibre synapse on CA3 pyramidal cells of rat hippocampus. *J. Physiol. (Lond.)* **472**, 615–663 (1993).
8. Spruston, N., Jonas, P. & Sakmann, B. Dendritic glutamate receptor channels in rat hippocampal CA3 and CA1 pyramidal neurons. *J. Physiol. (Lond.)* **482**, 325–352 (1995).
9. Major, G., Larkman, A. U., Jonas, P., Sakmann, B. & Jack, J. J. B. Detailed passive cable models of whole-cell recorded CA3 pyramidal neurons in rat hippocampal slices. *J. Neurosci.* **14**, 4613–4638 (1994).
10. Clay, J. R. & DeFelice, L. J. Relationship between membrane excitability and single channel open-close kinetics. *Biophys. J.* **42**, 151–157 (1983).
11. Clements, J. D., Lester, R. A. J., Tong, G., Jahr, C. E. & Westbrook, G. L. The time course of glutamate in the synaptic cleft. *Science* **258**, 1498–1501 (1992).
12. Stuart, G. J., Dodt, H.-U. & Sakmann, B. Patch-clamp recordings from the soma and dendrites of neurons in brain slices using infrared video microscopy. *Pflügers Arch.* **423**, 511–518 (1993).
13. Malinow, R. & Tsien, R. W. Presynaptic enhancement shown by whole-cell recordings of long-term potentiation in hippocampal slices. *Nature* **346**, 177–180 (1990).
14. Stevens, C. F. & Wang, Y. Changes in reliability of synaptic function as a mechanism for plasticity. *Nature* **371**, 704–707 (1994).
15. Isaac, J. T. R., Hjelmstad, G. O., Nicoll, R. A. & Malenka, R. C. Long-term potentiation at single fiber inputs to hippocampal CA1 pyramidal cells. *Proc. Natl Acad. Sci. USA* **93**, 8710–8715 (1996).
16. Allen, C. & Stevens, C. F. An evaluation of causes for unreliability of synaptic transmission. *Proc. Natl Acad. Sci. USA* **91**, 10380–10383 (1994).
17. Swanson, G. T., Kamboj, S. K. & Cull-Candy, S. G. Single-channel properties of recombinant AMPA receptors depend on RNA editing, splice variation, and subunit composition. *J. Neurosci.* **17**, 58–69 (1997).
18. Greengard, P., Jen, J., Nairn, A. C. & Stevens, C. F. Enhancement of the glutamate response by cAMP-dependent protein kinase in hippocampal neurones. *Science* **253**, 1135–1138 (1991).
19. Wang, L.-Y., Salter, M. W. & MacDonald, J. F. Regulation of kainate receptors by cAMP-dependent protein kinase and phosphatases. *Science* **253**, 1132–1135 (1991).
20. Cull-Candy, S. G. & Usowicz, M. M. Multiple-conductance channels activated by excitatory amino acids in cerebellar neurons. *Nature* **325**, 525–528 (1987).
21. Jahr, C. E. & Stevens, C. F. Glutamate activates multiple single channel conductances in hippocampal neurons. *Nature* **325**, 522–525 (1987).
22. Barria, A., Muller, D., Derkach, V., Griffith, L. C. & Soderling, T. R. Regulatory phosphorylation of AMPA-type glutamate receptors by CaM-KII during long-term potentiation. *Science* **276**, 2042–2045 (1997).
23. Kullmann, D. M. & Siegelbaum, S. A. The site of expression of NMDA receptor-dependent LTP: new fuel for old fire. *Neuron* **15**, 997–1002 (1995).
24. Shrike, A. M. & Malinow, R. Mechanisms of potentiation by calcium-calmodulin kinase II of postsynaptic sensitivity in rat hippocampal neurons. *J. Neurophysiol.* **78**, 2682–2692 (1997).
25. Liao, D., Hessler, N. A. & Malinow, R. Activation of postsynaptically silent synapses during pairing-induced LTP in CA1 region of hippocampal slice. *Nature* **375**, 400–404 (1995).
26. Isaac, J. T. R., Nicoll, R. A. & Malenka, R. C. Evidence for silent synapses: implications for the expression of LTP. *Neuron* **15**, 427–434 (1995).
27. Traynelis, S. F. & Wahl, P. Control of rat GluR6 glutamate receptor open probability by protein kinase A and calcineurin. *J. Physiol. (Lond.)* **503**, 513–531 (1997).
28. Hines, M. L. Efficient computation of branched nerve equations. *Int. J. Biomed. Comp.* **15**, 69–76 (1984).
29. Homes, W. R. & Levy, W. B. Insights into associative long-term potentiation from computational models of NMDA receptor-mediated calcium influx and intracellular calcium concentration changes. *J. Neurophysiol.* **63**, 1148–1168 (1990).
30. Sigworth, F. J. The variance of sodium current fluctuations at the node of Ranvier. *J. Physiol. (Lond.)* **307**, 97–129 (1980).

Acknowledgements. We thank A. Silver and M. Farrant for their advice, and W. W. Anderson for providing the data acquisition software. This work was supported by the MRC (T.A.B., G.L.C.), NIH (T.A.B.), Swiss National Science Foundation (A.L.) and Wellcome Trust (J.T.R.I., G.L.C.).

Correspondence and requests for materials should be addressed to G.L.C.

IgD can largely substitute for loss of IgM function in B cells

Claudia Lutz*, Birgit Ledermann†, Marie H. Kosco-Vilbois‡, Adrian F. Ochsenbein§, Rolf M. Zinkernagel§, Georges Köhler*¶ & Frank Brombacher*¶

* Max-Planck-Institute for Immunobiology, Freiburg, Germany

† Novartis Pharma Inc., Research, Basel, Switzerland

‡ Geneva Biomedical Research Institute, Glaxo Wellcome Research and Development, Plan-les-Quates, Switzerland

¶ Institute for Experimental Immunology, Zurich, Switzerland

§ University of Cape Town, Groote Schuur Hospital, Department of Immunology, H47, Observatory, 7925 Cape Town, South Africa

The μ and δ heavy chains of IgM and IgD, the first antibody isotypes expressed during bone-marrow B-cell development, are encoded by a common transcription unit. Expression of the μ chain on the surface of late pre-B cells allows their further development to immature B cells. Coexpression of the δ chain

¶ This work is dedicated to the memory of Georges Köhler who initiated this project.

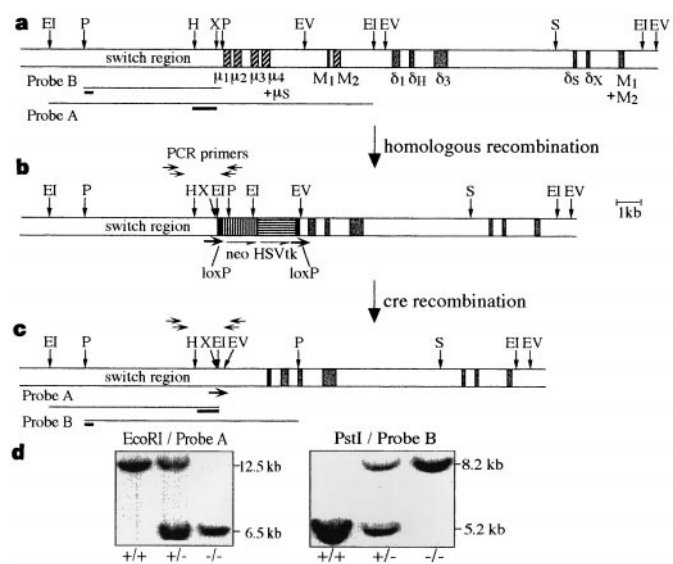


Figure 1 Generation of IgM-deficient mice. **a**, BALB/c $C\mu$ - $C\delta$ locus with switch region, exons (boxes), sites of restriction enzymes and length of fragments with probes used for Southern blotting. **b**, Mutated locus after homologous recombination in BALB/c ES cells with PCR primers (arrows) used for identification. **c**, Mutated locus after Cre-mediated site-specific recombination. **d**, Southern blot analysis of wild-type (+/+), heterozygous (+/-) and homozygous (-/-) mice. HSV tk, herpes simplex virus thymidine kinase; *neo*, neomycin-resistance gene; loxP, locus of crossing-over P1; restriction enzymes: El, *EcoRI*; EV, *EcoRV*; H, *HindIII*; P, *PstI*; S, *SalI*; X, *XbaI*.

and emigration of the immature B cells to the periphery eventually leads to the development of naive mature IgM/IgD double-positive cells. Although IgM is important in driving B-cell development¹, the contribution of IgD is not clear. Here we investigate the function of IgD. We generated mice deficient in IgM (IgM^{-/-} mice) by deleting the μ region in embryonic stem cells. IgM^{-/-} mice showed normal B-cell development and maturation, with IgD replacing membrane-bound and secretory IgM. Moreover, specific B-cell responses and isotype class switches occurred during immunization or infection. In contrast to mice deficient in B cells, IgM^{-/-} mice survived infection with vesicular stomatitis virus by developing neutralizing immunoglobulins, but they were more susceptible than wild-type controls with delayed specific immunoglobulin responses. These data lead us to conclude that IgD is largely able to substitute for IgM functions.

Generation of IgM^{-/-} mice by the deletion of the $C\mu$ and μ - δ intron, which contains the regulatory *cis*-elements for $C\delta$ expression^{2,3}, involved gene targeting⁴ and Cre-mediated site-specific recombination^{5,6} in BALB/c-derived embryonic stem (ES) cells⁷ (Fig. 1). The IgM^{-/-} mice showed no alterations in morphology or cell numbers in lymphatic organs compared to wild-type controls (data not shown). Fluorescence-activated cell sorting (FACS) analysis of bone-marrow cells from IgM^{-/-} mice revealed a normal number and distribution of B-cell receptor (BCR)-negative pro-B and pre-B cells as well as BCR-positive B cells (Fig. 2a, b). Furthermore, normal distribution of B220^{low}HSA^{high} immature and B220^{high}HSA^{low} mature recirculating B cells was observed in these mice (data not shown). IgM⁺ cells, however, were completely absent (Fig. 2b) and were replaced by IgD-positive cells in IgM^{-/-} mice (Fig. 2c). Instead, the numbers and distribution of peripheral mature IgD single-positive B cells were similar to those of IgM/IgD double-positive wild-type cells in spleen (Fig. 2d) and lymph nodes (data not shown). IgD intensity levels on B cells were increased in

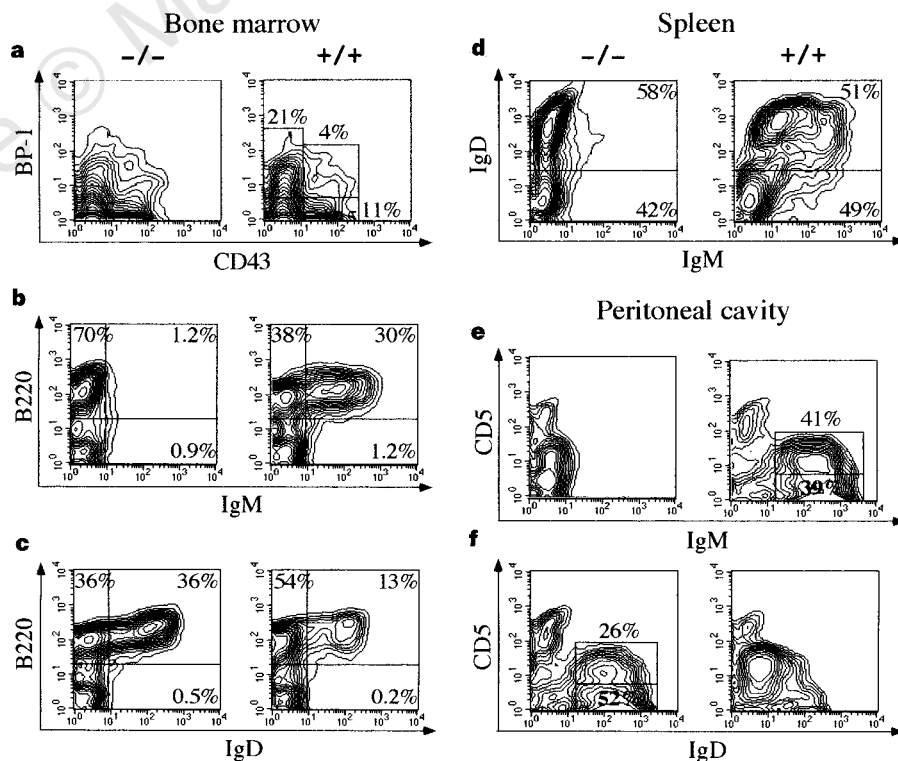


Figure 2 Flow cytometric analysis of IgM-deficient (-/-) or wild-type (+/+) mice. **a-c**, Bone-marrow cells. **a**, Only B220⁺ cells are shown; percentages of early (CD43⁺BP-1⁺), late pro- (CD43⁺BP-1⁺) and pre-B cells (CD43⁺BP-1⁺) are given. **b, c**, Only CD43⁺ cells are shown: IgM⁺ and/or IgD⁺ immature and mature B cells.

d, Spleen cells; IgM⁺ and/or IgD⁺ mature B cells. **e, f**, Peritoneal cavity cells. B-1a (Ig⁺CD5⁺) or B-1b sister and conventional B2 cells (Ig⁺CD5⁺) in boxes. A representative of 3 independent experiments with a total of 12 mice per group is shown. Significant differences were observed in the proportion of B-1a cells (see text).

IgM^{-/-} mice to a level comparable with wild-type IgM rather than IgD expression, as seen with peritoneal B cells (Fig. 2e, f). Consistent with the observed normal B-cell development and maturation in IgM^{-/-} mice, the lifespan of bone-marrow or splenic B cells were equivalent to wild-type B cells (data not shown).

Despite a normal conventional B-cell compartment, the proportion of self-replenishing peritoneal B220⁺CD5⁺ B-1a cells was reduced in IgM^{-/-} mice (26% ± 4.6%) when compared with wild-type mice (40% ± 6.9%; see Fig. 2e, f). B-1a cells are believed to produce natural autoantibodies, including anti-DNA-specific antibodies⁸ with a preference for λ light-chain⁹ expression. Consistently, serum λ light-chain-specific anti-DNA antibodies were significantly reduced in IgM^{-/-} mice compared with wild-type controls (1.79% ± 1.05% to 8.55% ± 2.97% of total λ-specific

antibodies of 5 mice per group), whereas κ light-chain-specific anti-DNA antibodies were present in the normal range (0.91% ± 0.20% to 1.32% ± 0.20% of total κ-specific antibodies). Autoantibodies were mainly of the IgD isotype in mutant and of the IgM isotype in wild-type mice (data not shown). A similar B-1a cell reduction was observed in mice deficient for the complement receptors CR1 and CR2 (ref. 10) or for CD19 (refs 11, 12). Taken together, these results suggest that antigen stimulation by the IgM B-cell receptor complex seems to favour maintenance of self-replenishing B-1a cells.

In contrast to high IgM (217 ± 65 μg ml⁻¹) but very low soluble IgD titres in wild-type mice (0.01 arbitrary units (AU)), no IgM (<10 pg ml⁻¹) but a striking increase of soluble IgD (sIgD) titre (121 ± 78 AU) was observed in IgM^{-/-} mice (6 individual mice per group analysed). Such an unusual IgD secretion was also found in plasmacytomas with a deleted μ gene¹³ or in mice transgenic for a rearranged δ gene¹⁴. These results strongly suggest that the deleted DNA fragment (Fig. 1) contains regulatory sequences suppressing sIgD expression. Alternatively, the deletion may have brought C_δ into the vicinity of (unknown) sequences that now enhance its secretion. The absence of IgM did not influence switching to other antibody isotypes in untreated mice or in mice infected with *Leishmania major* or *Nippostrongylus brasiliensis* (data not shown). After immunization with thymus-dependent (DNP-ova) or thymus-independent type II antigens (TNP-ficoll), normal hapten-specific antibody levels, isotype switches and kinetics were present, except for sIgM in wild-type mice which was replaced by sIgD in IgM^{-/-} mice (Fig. 3). Also, germinal centres were formed within IgD⁺ follicles in the spleen of immunized IgM^{-/-} mice (Fig. 4). The marginal zone B cells, normally IgM single-positive (Fig. 4a), were IgD single-positive in IgM^{-/-} mice (Fig. 4b, d). In addition, numerous IgD-producing plasma cells appeared to replace IgM producers present in wild-type mice (Fig. 4a, b, arrows). However, the localization of IgD in the IgM^{-/-} mice showed that the wild-type IgM was not totally substituted because immune complexes present on the follicular dendritic cell (FDC) network did not contain IgD (Fig. 4b, d). This observation would predict that FDC possess no receptors prone to trap IgD-containing immune complexes. Nevertheless, the FDC network was intact in the mutant mice (Fig. 4g, h) and retained IgG-containing immune complexes (Fig. 4i, j). Moreover, the T-cell zones were also unmodified as assessed by anti-Thy 1 labelling (data not

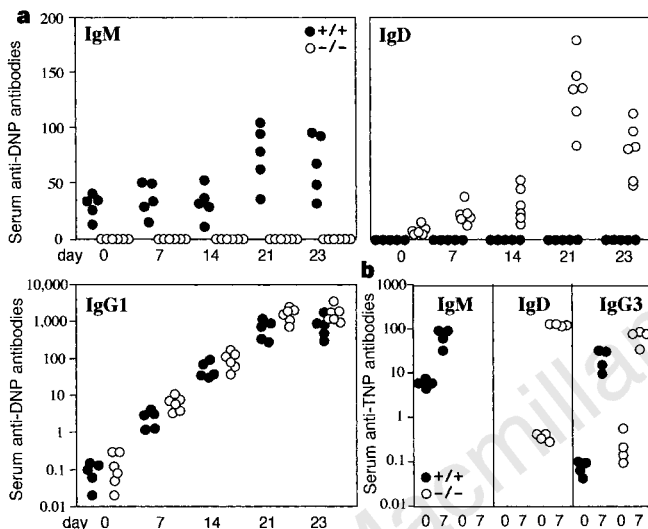


Figure 3 Immunization with antigen. Wild-type (filled circles) or IgM^{-/-} mice (open circles) were immunized with: **a**, thymus-dependent antigen DNP-ovalbumin at day 0 and rechallenged at day 14; or **b**, at day 0 with thymus-independent type II antigen TNP-ficoll. Production of anti-DNP- or -TNP-specific serum antibodies was determined as arbitrary units at the indicated days. Although pre-immune levels were nonspecifically higher in wild-type mice before immunization (**b**, day 0), the specific responses were equivalent after stimulation.

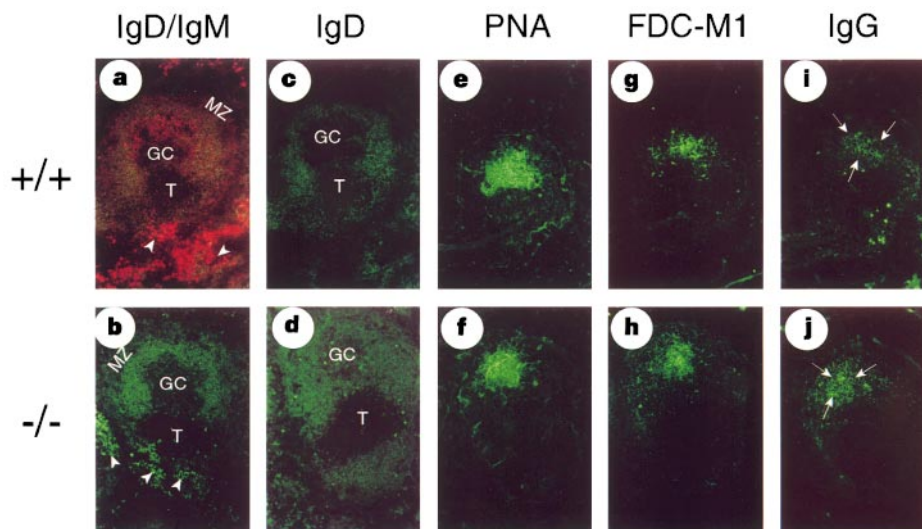


Figure 4 Germinal centre formation. Adjacent cryosections of wild-type (+/+) or IgM-deficient mice (-/-). **a, b**, Anti-IgD (green), anti-IgM (red) double staining; or **c, d**, anti-IgD staining only (green). GC, germinal centre; T, T-cell-rich paracortex; MZ, marginal zone. Plasma cells (arrows) produce IgM in wild-type or IgD in mutant

mice. **e, f**, PNA binding (green) to GC cells. **g, h**, Anti-FDC-M1 (green) staining to reveal the distribution of the follicular dendritic cell (FDC) network. **i, j**, Anti-IgG (green) staining showing trapping of IgG-containing immune complexes by FDCs (arrows) localized in the adjacent section (**g, h**), respectively.

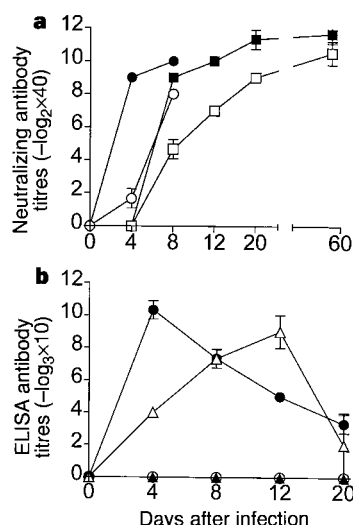


Figure 5 Immune responses to VSV in wild-type and IgM^{-/-} mice. Mice were infected with 10⁶ PFU VSV-IND at day 0. At the indicated time points post-infection: **a**, mean titres and standard deviation of neutralizing total immunoglobulin (circles) and IgG (squares) antibody titres; and **b**, VSV-specific IgM (circles) and IgD (triangles) were determined in wild-type (black symbols) or IgM^{-/-} mice (white symbols). Results of 2–3 individual mice are given. A representative of two independent experiments is shown.

shown). Taken together, many B-cell functions are normal in mice deficient in IgM.

Because of the structural difference of monomeric sIgD and pentameric or hexameric¹⁵ sIgM, the latter can induce the classic complement pathway and neutralize pathogens with high avidity. Vesicular stomatitis virus (VSV) belongs to the rabdoviridae and is closely related to the rabies virus. It is highly neurotropic and causes neurological disease and death in mice. Recovery from primary infection depends on neutralizing immunoglobulin antibody responses¹⁶. VSV induces rapid T-independent IgM responses¹⁷ and strong T-dependent¹⁸ neutralizing IgG responses in normal mice as shown after sublethal infection studies (Fig. 5a). Infected IgM^{-/-} mice produced very low neutralizing immunoglobulin titres around day 4 post-infection (Fig. 5a) with a marked increase up to day 8. At these time points only some of the total neutralizing immunoglobulins were of the IgG isotype. At 60 days post-infection, IgM^{-/-} mice reached antibody titres similar to those of wild-type mice (Fig. 5a). IgM specific for the neutralizing epitope of VSV (Fig. 5b) was found in control mice only, peaking at day 4 post-infection. On the other hand, VSV-specific IgD was found in IgM^{-/-} mice at day 4 post-infection (Fig. 5b) with highest serum titres at day 12. Therefore, the secretion of neutralizing IgD antibodies in IgM^{-/-} mice seems to compensate for the lack of IgM. The kinetics of antibody production, however, were markedly delayed compared to wild-type mice. During infections with cytopathic viruses, such as VSV, poliovirus and rabies virus, neutralizing antibodies have to be produced very soon after infection¹⁹, preventing dissemination of the virus to neuronal tissues. Mice deficient for B cells die during a VSV infection with 2 × 10⁶ plaque-forming units (PFU) of VSV-IND²⁰, whereas IgM^{-/-} mice survived this inoculum because of compensatory effects of IgD and the switch to neutralizing IgG antibodies. However, IgM^{-/-} mice showed increased susceptibility by a 10-fold decrease of the lethal VSV dose, from 2 × 10⁸ PFU in wild-type to 2 × 10⁷ PFU in IgM^{-/-} mice (all 4 mice per group died at day 8 and 9, respectively), probably a result of the delayed neutralizing antibody response in IgM^{-/-} mice.

Normal B-cell development in the absence of μ expression seems to conflict with a previously published mouse strain deficient for μ

expression (μ MT mice), which had a complete block at the pre-B cell stage²¹. In μ MT mice the heavy-chain locus was disrupted by insertion of a stop codon and a *neo*^r-gene into the first membrane exon (μ M1) of C μ , and neither IgM nor IgD surface expression was found in these mice²¹. Apparently these mutations abrogated both μ and δ expression. Mice deficient for IgD also have no developmental defect^{22,23}. Therefore, it seems that early expression of either μ or δ is sufficient to promote B-cell development. The formation of a pre-B-cell receptor, a crucial step in the further development of B cells²⁴, seems possible with the δ as well as the μ heavy chain. This may be due to both chains having a similar short cytoplasmic tail and the ability of Ig α and Ig β , the signal-transducing unit of the BCR, to associate with all immunoglobulin classes²⁵. In summary, our studies using IgM^{-/-} mice showed that IgD can largely substitute for IgM in driving B-cell development, maturation and function, suggesting redundancy between these isotypes. Furthermore, the IgM-deficient mouse line has proved to be an ideal animal model in which to study the consequences of loss of IgM functions. □

Methods

Gene targeting techniques. The targeting construct consisted of the 0.9 kilobase *HindIII*–*XbaI* fragment (short arm) 5' of exon μ 1 and the *EcoRV*–*Sall* fragment (long arm) containing exon δ 1, δ H and δ 3 of genomic BALB/c-derived DNA. The selection cassette, consisting of *neo*^r and HSV, is flanked by loxP-sites²⁶ and was introduced by a *Sall* linker. BALB/c ES cells⁷ were grown as described²⁷, transfected with 25 μ g linearized targeting vector by electroporation for homologous recombination and selected with G418-containing medium. For site-specific recombination, targeted ES cells were transiently transfected with 20 μ g of supercoiled pMC-CreN (a gift from H. Gu and modified by L. Nitschke) and selected for ganciclovir resistance. Selected ES cell clones were picked and analysed by nested primer polymerase chain reaction (PCR) analysis as described²⁸. Correct recombination was confirmed by Southern blot analysis. Successfully targeted ES cells were injected in C57BL/6-derived blastocysts (3.5 days post-coitus) and reimplanted in the uteri of 2.5-day pseudopregnant females. Germline transmission was obtained and mice were bred to homozygosity.

Flow cytometric analysis. Analysis of bone marrow, spleen and peritoneal exudate cells were carried out on a FACScan flow cytometer (Becton Dickinson, Heidelberg, Germany) as described²⁹. The antibodies used for staining (anti-B220/CD45R (RA3-6B2); anti-CD43 (S7); anti-BP-1 (6C3); anti-IgM^d (RS3.1); anti-IgD^d (AMS15.1.5); anti-CD5 (53-7.3); anti-HSA (CD24, M1/69)) were either biotinylated and counterstained with streptavidin-red670 (Gibco, Scotland), or directly fluoresceine isothiocyanate (FIC)- or phycoerythrin-conjugated (PharMingen, San Diego, USA).

Enzyme-linked immunosorbant assay (ELISA) and immunizations. (1) Immunoglobulin isotype levels of serial-diluted sera from 6.5-week-old untreated mice were determined by ELISA using class-specific unlabelled, alkaline phosphatase (AP)-labelled and standard antibodies (Southern Biotechnology). (2) Anti-DNA-antibodies were bound on ELISA plates coated with 10 μ g ml⁻¹ of double-stranded salmon sperm DNA and detected with anti- κ or anti- λ AP-labelled antibodies (Southern Biotechnology). (3) For 2,4-dinitrophenyl-ovalbumin (DNP-ova) or 2,4,6-trinitrophenyl (TNP)-ficoll immunization, mice 8–10 weeks old were injected subcutaneously (s.c.) at day 0 with 20 μ g of alum-precipitated DNP-ova and *Bordetella pertussis* as adjuvants (Serva, Heidelberg) or intraperitoneally (i.p.) with 10 μ g TNP-ficoll. DNP-ova immunized mice were rechallenged at day 14 s.c. and i.p. with 20 μ g DNP-ova. DNP- or TNP-specific antibodies in the sera were bound by hapten-BSA-coated plates and detected by anti-mouse isotype-specific antibodies as described above. (4) The VSV-specific ELISA has been described previously²⁰.

Immunohistochemistry. Cryosections were prepared from spleens of DNP-ova immunized IgM^{-/-} and wild-type mice (day 9 after secondary immunization). The sections were labelled for IgM (Texas Red conjugated, Southern Biotechnology), IgD (FITC conjugated, PharMingen), PNA³⁰ (biotinylated, Vector Laboratories) followed by streptavidin-FITC (Sigma), FDC-M1 (rat anti-mouse FDC, revealed by Fab₂ fragments of mouse anti-rat IgG-specific FITC, Jackson Immunoresearch Labs) or IgG1 (ATC conjugated, Southern Biotechnology).

VSV infection and serum neutralization test. After immunization with 2×10^6 PFU VSV-IND i.v. serum was collected from mice at specified times, prediluted 40-fold in MEM containing 5% FCS and heat inactivated for 30 min at 56 °C. Serial twofold dilutions were mixed with equal volumes of VSV (500 PFU ml⁻¹) and incubated for 90 min at 37 °C in an atmosphere with 5% CO₂. Then 100 µl of the serum-virus mixture was transferred onto Vero cell monolayers in 96-well plates and incubated for 1 h at 37 °C. The monolayers were then overlaid with 100 µl of DMEM containing 1% methylcellulose and incubated for 24 h at 37 °C. The overlay was flicked off and the monolayer was fixed and stained with 0.5% crystal violet. The highest dilution of serum that reduced the number of plaques by 50% was taken as the titre. To determine IgG titres, undiluted serum was pretreated with an equal volume of 0.1 mM β-2-mercaptoethanol in saline.

Received 19 December 1997; accepted 16 April 1998.

- Iglesias, A. *et al.* Early B cell development requires mu signalling. *Eur. J. Immunol.* **23**, 2622–2630 (1993).
- Ashfield, R. *et al.* MAZ-dependent termination between closely spaced human complement genes. *EMBO J.* **13**, 5656–5667 (1994).
- Yuan, D., Witte, P. L., Tan, J., Hawley, J. & Dang, T. Regulation of IgM and IgD heavy chain gene expression: effect of abrogation of intergenic transcriptional termination. *J. Immunol.* **157**, 2073–2081 (1996).
- Mansour, S. L., Thomas, K. R. & Capecchi, M. R. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* **336**, 348–352 (1988).
- Sauer, B. Manipulation of transgenes by site-specific recombination: use of Cre recombinase. *Methods Enzymol.* **225**, 890–900 (1993).
- Gu, H., Zou, Y. R. & Rajewsky, K. Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* **73**, 1155–1164 (1993).
- Noben-Trauth, N., Kohler, G., Burki, K. & Ledermann, B. Efficient targeting of the IL-4 gene in a BALB/c embryonic stem cell line. *Trans. Res.* **5**, 487–491 (1996).
- Hayakawa, K. *et al.* Ly-1 B cells: functionally distinct lymphocytes that secrete IgM autoantibodies. *Proc. Natl Acad. Sci. USA* **81**, 2494–2498 (1984).
- Hayakawa, K., Hardy, R. R. & Herzenberg, L. A. Peritoneal Ly-1 B cells: genetic control, autoantibody production, increased lambda light chain expression. *Eur. J. Immunol.* **16**, 450–456 (1986).
- Ahearn, J. M. *et al.* Disruption of the Cr2 locus results in a reduction in B-1a cells and in an impaired B cell response to T-dependent antigen. *Immunity* **4**, 251–262 (1996).
- Engel, P. *et al.* Abnormal B lymphocyte development, activation, and differentiation in mice that lack or overexpress the CD19 signal transduction molecule. *Immunity* **3**, 39–50 (1995).
- Rickert, R. C., Rajewsky, K. & Roes, J. Impairment of T-cell-dependent B cell responses and B-1 cell development in CD19-deficient mice. *Nature* **376**, 352–355 (1995).
- Weiss, E. A., Tucker, P. W., Finkelman, F. D. & Yuan, D. Analysis of immunoglobulin heavy chain delta transcription termination in the production of delta S or delta M mRNA. *Mol. Immunol.* **28**, 687–695 (1991).
- Iglesias, A., Lamers, M. & Kohler, G. Expression of immunoglobulin delta chain causes allelic exclusion in transgenic mice. *Nature* **330**, 482–484 (1987).
- Randall, T. D., King, L. B. & Corley, R. B. The biological effects of IgM hexamer formation. *Eur. J. Immunol.* **20**, 1971–1979 (1990).
- Bachmann, M. F. *et al.* The role of antibody concentration and avidity in antiviral protection. *Science* **276**, 2024–2027 (1997).
- Freer, G. *et al.* Vesicular stomatitis virus Indiana glycoprotein as a T-cell-dependent and -independent antigen. *J. Virol.* **68**, 3650–3655 (1994).
- Leist, T. P., Cobbold, S. P., Waldmann, H., Aguet, M. & Zinkernagel, R. M. Functional analysis of T lymphocyte subsets in antiviral host defense. *J. Immunol.* **138**, 2278–2281 (1987).
- Kalinka, U. *et al.* Monovalent single-chain Fv fragments and bivalent miniantibodies bound to vesicular stomatitis virus project against lethal infection. *Eur. J. Immunol.* **26**, 2801–2806 (1996).
- Brundler, M. A. *et al.* Immunity to viruses in B cell-deficient mice; influence of antibodies on virus persistence and on T cell memory. *Eur. J. Immunol.* **26**, 2257–2262 (1996).
- Kitamura, D., Roes, J., Kuhn, R. & Rajewsky, K. A B cell-deficient mouse by targeted disruption of the membrane exon of the immunoglobulin µ chain gene. *Nature* **350**, 423–426 (1991).
- Nitschke, L., Kosco, M. H., Kohler, G. & Lamers, M. C. Immunoglobulin D-deficient mice can mount normal immune responses to thymus-independent and -dependent antigens. *Proc. Natl Acad. Sci. USA* **90**, 1887–1891 (1993).
- Roes, J. & Rajewsky, K. Immunoglobulin D (IgD)-deficient mice reveal an auxiliary receptor function for IgD in antigen-mediated recruitment of B cells. *J. Exp. Med.* **177**, 45–55 (1993).
- Shinkai, Y. *et al.* RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* **68**, 855–867 (1992).
- Reth, M. Antigen receptors on B lymphocytes. *Annu. Rev. Immunol.* **10**, 97–121 (1992).
- Achatz, G., Nitschke, L. & Lamers, M. C. Effect of transmembrane and cytoplasmic domains of IgE on the IgE response. *Science* **276**, 409–411 (1997).
- Mummery, C. L., Feyen, A., Freund, E. & Shen, S. Characteristics of embryonic stem cell differentiation: a comparison with two embryonal carcinoma cell lines. *Cell Diff. Dev.* **30**, 195–206 (1990).
- Nitschke, L., Kopf, M. & Lamers, M. C. Quick nested PCR screening of ES cell clones for gene targeting events. *Biotechnology* **14**, 914–916 (1993).
- Brombacher, F., Kohler, G. & Eibell, H. B cell tolerance in mice transgenic for anti-CD8 immunoglobulin µ chain. *J. Exp. Med.* **174**, 1335–1346 (1991).
- Coico, R. F., Bhogal, B. S. & Thorbecke, G. J. Relationship of germinal centers in lymphoid tissue to immunologic memory. VI. Transfer of B cell memory with lymph node cells fractionated according to their receptors for peanut agglutinin. *J. Immunol.* **131**, 2254–2257 (1983).

Acknowledgements. We thank M. Held, C. Westphal, I. Fidler, S. Meier, S. Herren and K.-H. Widmann for technical help, and L. Nitschke, P. Nielsen and M. Reth for discussion and critically reviewing the manuscript.

Correspondence and requests for materials should be addressed to F.B. (e-mail: fbrombac@uctgsh1.uct.ac.za).

Evidence for the shikimate pathway in apicomplexan parasites

Fiona Roberts^{††‡}, Craig W. Roberts^{††§}, Jennifer J. Johnson[‡], Dennis E. Kyle^{||}, Tino Krell^{||}, John R. Coggins[†], Graham H. Coombs[†], Wilbur K. Milhous^{||}, Saul Tzipori[#], David J. P. Ferguson[☆], Debopam Chakrabarti^{**} & Rima McLeod[‡]

[‡] Michael Reese Hospital and Medical Center, Chicago, and The University of Chicago, Chicago, Illinois 60616, USA

[§] Department of Immunology, University of Strathclyde, Glasgow G4 0NR, UK
^{||} Walter Reed Army Institute of Research, Washington DC 20307-5100, USA

[†] Institute of Biomedical & Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK

[#] Tufts University School of Veterinary Medicine, Boston, Massachusetts 01536, USA

[☆] John Radcliffe Hospital, Oxford University, Oxford OX3 9DU, UK

^{**} University of Central Florida, Orlando, Florida 32816-2360, USA

[†] These authors contributed equally to this work.

Parasites of the phylum Apicomplexa cause substantial morbidity, mortality and economic losses, and new medicines to treat them are needed urgently^{1,2}. The shikimate pathway is an attractive target for herbicides and antimicrobial agents because it is essential in algae, higher plants, bacteria and fungi, but absent from mammals^{3,4}. Here we present biochemical, genetic and chemotherapeutic evidence for the presence of enzymes of the shikimate pathway in apicomplexan parasites. *In vitro* growth of *Toxoplasma gondii*, *Plasmodium falciparum* (malaria) and *Cryptosporidium parvum* was inhibited by the herbicide glyphosate, a well-characterized inhibitor³ of the shikimate pathway enzyme 5-enolpyruvyl shikimate 3-phosphate synthase. This effect on *T. gondii* and *P. falciparum* was reversed by treatment with *p*-aminobenzoate, which suggests that the shikimate pathway supplies folate precursors for their growth. Glyphosate in combination with pyrimethamine limited *T. gondii* infection in mice. Four shikimate pathway enzymes were detected in extracts of *T. gondii* and glyphosate inhibited 5-enolpyruvyl shikimate 3-phosphate synthase activity. Genes encoding chorismate synthase, the final shikimate pathway enzyme, were cloned from *T. gondii* and *P. falciparum*. This discovery of a functional shikimate pathway in apicomplexan parasites provides several targets for the development of new antiparasite agents.

We hypothesized that the plastid-like organelle of apicomplexans^{1,2,5–7} may have essential biosynthetic functions similar to those of plant plastids. We found that *T. gondii* bradyzoites have a plastid organelle morphologically similar (data not shown) to that of tachyzoites^{2,6}. This finding suggested that the plastid of this untreatable bradyzoite life-cycle stage, as well as tachyzoites, might contain new plastid-associated antimicrobial agent targets. Plant plastids are the site for many essential biochemical pathways including the biosynthesis of folate, amino acids, ubiquinone, haem, nucleotide, lipid and starch. Plant plastid organelles contain the shikimate pathway in which erythrose 4-phosphate and phosphoenolpyruvate are converted to chorismate in seven enzymatic steps^{4,8}. All seven enzymes are encoded in the nucleus, made in the cytoplasm and targeted to the chloroplast^{9,10}. In plants, chorismate is an essential substrate for the synthesis of *p*-aminobenzoate (PABA) and folate and is also required for the synthesis of ubiquinone, aromatic amino acids and almost all other aromatic

[†] Present address: University of Strathclyde, Glasgow, UK (C.W.R.); University of Glasgow, Glasgow, UK (E.R.).

compounds⁴. Glyphosate (*N*-(phosphonomethyl) glycine) is a highly effective herbicide because of its potent and specific inhibition of 5-enolpyruvyl shikimate 3-phosphate (EPSP) synthase, the sixth enzyme of the shikimate pathway^{3,4}. Mammals and many protists rely exclusively on exogenous folates and do not have a shikimate pathway; its presence in apicomplexan parasites would make it an excellent target for new anti-parasite drugs.

To investigate whether the shikimate pathway is present in *T. gondii* (Fig. 1), we first determined whether glyphosate¹¹ affects *T. gondii*. Glyphosate inhibited the growth of *T. gondii* tachyzoites *in vitro* (Fig. 1a); glyphosate inhibition of *T. gondii* was demonstrated microscopically and as diminished uptake of uracil by the parasite (3.12 mM glyphosate versus none, $P < 0.002$). The herbicide at this concentration had no effect on the growth of mammalian cells. By using similar methodology, we found that glyphosate at a concentration of 12.5 mM inhibited growth of two different cell lines (HFF and MBDK) by 30%, there being no effect at all at 9 mM. These findings suggested a specific action of glyphosate against the parasite and thus the presence of a shikimate pathway, although confirmatory data were necessary. To this end, we looked for antagonists of the anti parasite action. The effect was reversed by the addition of PABA (Fig. 1b). In parallel control experiments, PABA abolished the growth inhibitory effect of sulphadiazine, a well-characterized inhibitor of PABA incorporation and folate biosynthesis, but not pyrimethamine, which acts later in the pathway at the dihydrofolate reductase step (Fig. 1b). To confirm that EPSP synthase was indeed the target of glyphosate action, the enzyme activity was assayed in crude *T. gondii* extracts (0.34 mU mg⁻¹ protein; Fig. 1c) by measuring directly the formation

of EPSP using a method based on high-performance liquid chromatography (HPLC). The identity of the product was confirmed to be EPSP by comparing its retention time and ultraviolet spectrum with authentic samples of EPSP generated by biotransformation¹². Glyphosate (1 mM) inhibited the enzyme activity by 85% (Fig. 1c).

These data establish the presence of EPSP synthase activity in *T. gondii* and suggest that there is a metabolic route for the synthesis of PABA and folate from EPSP. We also detected chorismate synthase, the enzyme following EPSP synthase in the shikimate pathway, at low levels using an anaerobic assay¹³. Genetic evidence for the presence of this enzyme in the parasite was also sought. A partial complementary DNA sequence of about 300 bases was identified from the *T. gondii* project (EST TgEST zyllc05.r1; D. Sibley *et al.*, Washington University Parasite Genome Blast Server: www.ebi.ac.uk/parasites/parasite_blast_server.html). This sequence, when translated, had about 30% homology with chorismate synthase from several organisms. Both strands of the corresponding clone were sequenced and found to be 2,312 bases in length (Genbank accession number U 93689). Analysis revealed a long open reading frame (ORF) of 1,608 base pairs which encodes a 536 amino-acid protein. The deduced amino-acid sequence has considerable identity with chorismate synthases of several species (Fig. 2): *Plasmodium falciparum* (49.7%), *Synechocystis* (51.4%), *Saccharomyces cerevisiae* (49.6%), *Solanum lycopersicum* (47.2%), *Neurospora crassa* (45.0%) and *Haemophilus influenza* (44.5%). The large inserts in the apicomplexan sequences that have no counterparts in the genes of other organisms were not included in these calculations. The putative *T. gondii* protein differs from other known chorismate synthases in length. It shares all of the amino

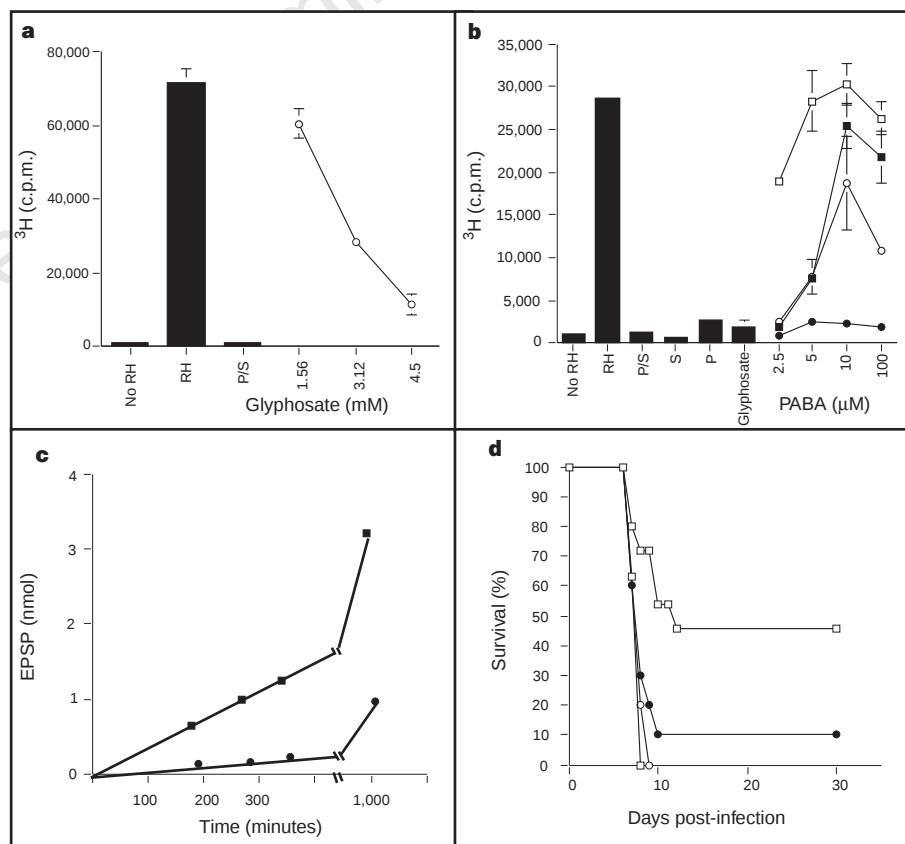


Figure 1 Functional and enzymatic evidence for the shikimate pathway in *T. gondii*. **a**, Glyphosate inhibition of *T. gondii* tachyzoites growth in human foreskin fibroblasts *in vitro*. RH, *T. gondii* strain; P/S, pyrimethamine and sulphadiazine. **b**, PABA rescue of glyphosate growth inhibition. PABA, *para*-amino benzoic acid; S, sulphadiazine; Pyr, pyrimethamine. **c**, Inhibition of EPSP synthase enzyme

activity by 1 mM glyphosate. Squares, without glyphosate. Circles, with glyphosate. **d**, *In vitro* activity against *T. gondii* of glyphosate (100 mg kg⁻¹ day⁻¹) (open squares) and pyrimethamine (12.5 mg kg⁻¹ day⁻¹) (filled squares), alone or in combination (open circles); controls (filled circles).

acids known to be highly conserved in chorismate synthases from other organisms but has a number of insertions, which, on the basis of their amino-acid composition, appear to form exposed loops. We have also cloned an ORF from *P. falciparum* (Genbank accession number AF008549). This also has unusual insertions, suggesting

that these may be a feature of apicomplexan chorismate synthases. The function(s) of the insertions that have no counterparts in the genes of other organisms remain to be determined.

Glyphosate also was shown to inhibit the growth *in vitro* of the asexual erythrocytic forms of *P. falciparum* (Fig. 3a). Glyphosate (1.08 mM) completely inhibited *in vitro* growth of asexual stages of the D6 clone. Doxycycline and proguanil, medications currently used in combination for treatment of *P. falciparum* malaria, have 50% and 90% inhibitory concentrations (IC_{50} and IC_{90}) *in vitro* of the same magnitude as glyphosate. Both PABA and folic acid antagonized the inhibitory effect of glyphosate. PABA was about 1,000-fold more effective than folic acid. The growth *in vitro* of *C. parvum*, another opportunistic apicomplexan parasite of immunosuppressed individuals, was similarly found to be reduced by glyphosate (Fig. 3b, $P < 0.001$). However, in this case PABA did not appear to ablate the inhibition. This may be a reflection of the poor transport of PABA into the parasite. Alternatively, in *C. parvum*, glyphosate inhibition of the shikimate pathway could have a more crucial effect on other pathways downstream from chorismate.

These data taken together provide strong evidence that the pathway linking shikimate phosphate (the substrate of EPSP synthase) and folate is present in apicomplexan parasites. We have also obtained evidence for the presence of two of the earlier enzymes in the shikimate pathway. The third enzyme, 3-dehydroquinate dehydratase, was detected in *T. gondii* lysates (3.0 mU per mg protein) by direct spectrophotometric assay¹². A partially purified enzyme (21 mU per mg protein) was shown to be heat labile, indicating that it is similar to the type I dehydroquinate dehydratases

Synechocystis	-----MGNITFG
S. lycopersicum	MASFVPTKQFVGASSSSDIGSSRLVSLQLPSKSSNFHLPSPRSQKRLKLEIQAGSTFG
S. cerevisiae	-----MSTFG
N. crassa	-----MSTFG
H. influenza	-----MAGNTIG
T. gondii	-----MSSYG
P. falciparum	-----MSTYG
Synechocystis	SLFRITTFGESHGGGVGIIIDGCPPLRLSPESQVLDLRRRPGQSKITTPRKEADQCEI
S. lycopersicum	NYFRVTTGESHGGGVGIIIDGCPPLRLSPESQVLDLRRRPGQSKITTPRKEADQCEI
S. cerevisiae	KLFRVTTGESHGCKSVGCIIDGCPPLRLSPESQVLDLRRRPGQSKITTPRKEADQCEI
N. crassa	HYFRVTTGESHGCKSVGCIIDGCPPLRLSPESQVLDLRRRPGQSKITTPRKEADQCEI
H. influenza	QLFRVTTGESHGCKSVGCIIDGCPPLRLSPESQVLDLRRRPGQSKITTPRKEADQCEI
T. gondii	AALRHITFGESHGCKSVGCIIDGCPPLRLSPESQVLDLRRRPGQSKITTPRKEADQCEI
P. falciparum	TLKVTSYGESHGCKSVGCIIDGCPPLRLSPESQVLDLRRRPGQSKITTPRKEADQCEI
Synechocystis	LSGVFEGKTLGTPAILVRNKDARSQDYN--EMAVKYPFSHADATYEAQYGRNQGGGR
S. lycopersicum	SSGTADGLTIGSPKRVFVPTDQNDYS--EMSLAYRFSHADATYDFKYGVRVSGQGGGR
S. cerevisiae	QSTFEGKTLGTPAILVRNKDARSQDYN--EMAVKYPFSHADATYEAQYGRNQGGGR
N. crassa	LSGVFEGKTLGTPAILVRNKDARSQDYN--EMAVKYPFSHADATYEAQYGRNQGGGR
H. influenza	LSGVFEGKTLGTPAILVRNKDARSQDYN--EMAVKYPFSHADATYEAQYGRNQGGGR
T. gondii	LSGVFEGKTLGTPAILVRNKDARSQDYN--EMAVKYPFSHADATYEAQYGRNQGGGR
P. falciparum	LSGVFEGKTLGTPAILVRNKDARSQDYN--EMAVKYPFSHADATYEAQYGRNQGGGR
Synechocystis	SSARETIGRVAAGAIKAKLQAQNGVEIVAVKSIQDI
S. lycopersicum	SSARETIGRVAAGAIKAKLQAQNGVEIVAVKSIQDI
S. cerevisiae	ASARETIGRVAAGAIKAKLQAQNGVEIVAVKSIQDI
N. crassa	SSARETIGRVAAGAIKAKLQAQNGVEIVAVKSIQDI
H. influenza	SSARETIGRVAAGAIKAKLQAQNGVEIVAVKSIQDI
T. gondii	SSARETIGRVAAGAIKAKLQAQNGVEIVAVKSIQDI
P. falciparum	SSARETIGRVAAGAIKAKLQAQNGVEIVAVKSIQDI
Synechocystis	-----EATVDNNTVLE
S. lycopersicum	-----EDLIDHNTVLE
S. cerevisiae	-----RDSFDPEFQHLNLTITRE
N. crassa	-----SPSTNPEFQHLNLTITRE
H. influenza	-----OK-VGGIDWE
T. gondii	RLGVVRVSPDGTTFDLNANRLDYE
P. falciparum	SYGTVRNEKEKIFMDCFNRIYDMNASLKTDEYNKNTLTPISDNTYINVKNTMNCIN
Synechocystis	-----IVRCPDECAEKMIERIDQVLAQKDSIGGVV
S. lycopersicum	-----IVRCPDECAEKMIERIDQVLAQKDSIGGVV
S. cerevisiae	-----IVRCPDECAEKMIERIDQVLAQKDSIGGVV
N. crassa	-----IVRCPDECAEKMIERIDQVLAQKDSIGGVV
H. influenza	-----IVRCPDECAEKMIERIDQVLAQKDSIGGVV
T. gondii	-----IVRCPDECAEKMIERIDQVLAQKDSIGGVV
P. falciparum	-----IVRCPDECAEKMIERIDQVLAQKDSIGGVV
Synechocystis	ECAIRNAPKGLGEPVFDKLEADLAKAMSLPATKGFEGSGFAGTLTGSQHNDDEYLYDE
S. lycopersicum	TCIVRNLPGLGTPVFDKLEADLAKAMSLPATKGFEGSGFAGTLTGSQHNDDEYLYDE
S. cerevisiae	TCIVRNLPGLGTPVFDKLEADLAKAMSLPATKGFEGSGFAGTLTGSQHNDDEYLYDE
N. crassa	TCIVRNLPGLGTPVFDKLEADLAKAMSLPATKGFEGSGFAGTLTGSQHNDDEYLYDE
H. influenza	TCIVRNLPGLGTPVFDKLEADLAKAMSLPATKGFEGSGFAGTLTGSQHNDDEYLYDE
T. gondii	TCIVRNLPGLGTPVFDKLEADLAKAMSLPATKGFEGSGFAGTLTGSQHNDDEYLYDE
P. falciparum	TCIVRNLPGLGTPVFDKLEADLAKAMSLPATKGFEGSGFAGTLTGSQHNDDEYLYDE
Synechocystis	-----AGE
S. lycopersicum	-----HGR
S. cerevisiae	-----ETN
N. crassa	-----NTEIPPSVAASGAARNGI
H. influenza	-----FES
T. gondii	-----RASCSFSEASATTKHERDGSATLSREASDGRITTSHEEVEVERGRERIQDRLHVTG
P. falciparum	NMSTKESDLYDDKGECKNMSYHSTIQNEDQILNSTKGFMPKNDKFNINDDYNTF
Synechocystis	-----WRTRINRSQGVQGG
S. lycopersicum	-----IRTRINRSQGVQGG
S. cerevisiae	-----RLRTKTNNSQGVQGG
N. crassa	-----PRPKLTITKTNNSQGVQGG
H. influenza	-----NHAGGILGG
T. gondii	-----VDQNGNSDESVRTSKSEASITRLSGNAASGAPVCRIPLGEGVIRCGSNAGGTLAG
P. falciparum	-----NNN-----EKKLLITKTNNSQGVQGG
Synechocystis	ISNGEPIIMRIAFKPTATIGQEQKTVSNIG--EETTLAAGRHDPVCPVRAVPMVEAMAL
S. lycopersicum	ISNGEPIIMRIAFKPTATIGQEQKTVSNIG--EETTLAAGRHDPVCPVRAVPMVEAMAL
S. cerevisiae	ISNGEPIIMRIAFKPTATIGQEQKTVSNIG--EETTLAAGRHDPVCPVRAVPMVEAMAL
N. crassa	ISNGEPIIMRIAFKPTATIGQEQKTVSNIG--EETTLAAGRHDPVCPVRAVPMVEAMAL
H. influenza	ISNGEPIIMRIAFKPTATIGQEQKTVSNIG--EETTLAAGRHDPVCPVRAVPMVEAMAL
T. gondii	ISNGEPIIMRIAFKPTATIGQEQKTVSNIG--EETTLAAGRHDPVCPVRAVPMVEAMAL
P. falciparum	ISNGEPIIMRIAFKPTATIGQEQKTVSNIG--EETTLAAGRHDPVCPVRAVPMVEAMAL
Synechocystis	VLCDHLLRFQAGCKTL
S. lycopersicum	VLCDHLLRFQAGCKTL
S. cerevisiae	VLCDHLLRFQAGCKTL
N. crassa	VLCDHLLRFQAGCKTL
H. influenza	VLCDHLLRFQAGCKTL
T. gondii	VLCDHLLRFQAGCKTL
P. falciparum	VLCDHLLRFQAGCKTL

Figure 2 Comparison of deduced amino-acid sequences of *T. gondii*, *P. falciparum* and other organisms' chorismate synthases. Amino acids that are identical in all seven organisms are indicated by *; : indicates conservative substitutions among all sequences; . indicates less conservative substitutions among all sequences as defined in the clustal program. Preliminary homology searches using the BLAST program show a small degree of homology of the *T. gondii* chorismate synthase insertion with a number of unrelated proteins including the amyloplast-specific transit peptide of the wx+ protein (glucosyl transferase) from *Zea mays*.

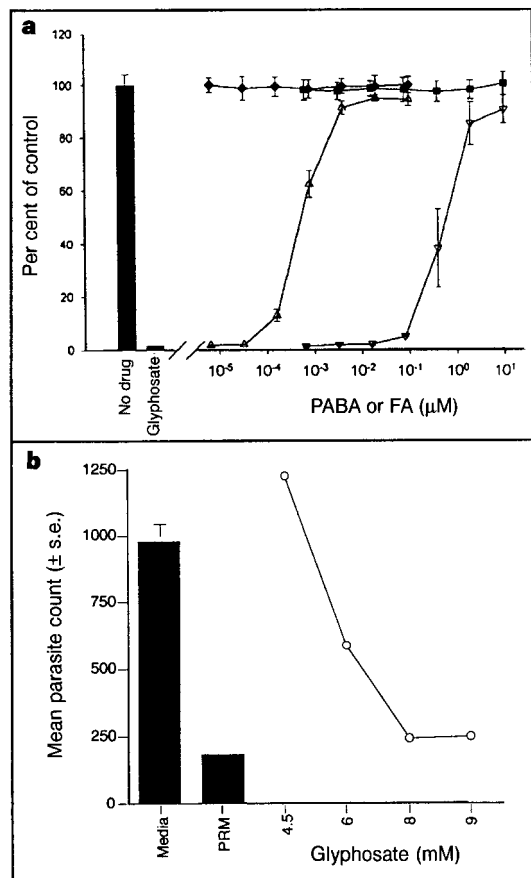


Figure 3 Evidence for the shikimate pathway in *P. falciparum* and *C. parvum*. **a**, Functional evidence for the shikimate pathway in *P. falciparum*. Glyphosate inhibition of *in vitro* growth of asexual erythrocytic forms and PABA and folic acid antagonism of growth inhibition. **b**, Glyphosate inhibition of *C. parvum* growth. PRM, paromomycin.

Table 1 Effects of glyphosate alone and together with antimicrobial agents that inhibit folate synthesis on replication of *T. gondii*

Drug A	Drug B	Untreated (c.p.m.)	A (c.p.m.)	B (c.p.m.)	A + B (c.p.m.)		Ratio (actual: predicted)*
					Actual	Predicted	
Glyphosate	Sulphadiazine	71,449 ± 3,763	28,138 ± 2,216	25,026 ± 4,365	2,368 ± 418	9,856	0.24
Glyphosate	Pyrimethamine	64,343 ± 1,222	25,097 ± 1,398	69,217 ± 3,253	9,354 ± 2126	25,097	0.37

Predicted c.p.m. of ^3H for drug A + drug B (if effect is only additive, not synergistic) is calculated as (c.p.m. drug A × c.p.m. drug B) divided by c.p.m. of untreated culture. Concentrations were: glyphosate (3.125 mM); sulphadiazine (6.25 $\mu\text{g ml}^{-1}$); pyrimethamine (0.025 $\mu\text{g ml}^{-1}$).

* A ratio of actual: predicted of <1 is considered synergistic. A ratio of actual: predicted ≥1 is considered additive.

found in higher plants¹⁴. The fifth enzyme of the pathway (shikimate kinase) was detected in *T. gondii* lysates (1.6 mU per mg protein) by measuring directly the formation of shikimate 3-phosphate from shikimate using an HPLC-based assay¹⁵. The identity of the product was confirmed by observing its conversion to EPSP after the addition of phosphoenol pyruvate and purified *Escherichia coli* EPSP synthase, as described previously¹⁶. Shikimate kinase, 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase and shikimate dehydratase activities previously had been reported at very low levels in crude extracts of *P. falciparum*¹⁷.

The partial inhibition of growth *in vitro* of the three parasites by glyphosate, together with the reversal of the effect for two of these parasites by PABA or folate, suggests that the shikimate pathway is important at least for the supply of folate in these organisms. The possibility that glyphosate could act in conjunction with conventional antifolate medicines was therefore tested. Using *in vitro* assays, glyphosate was found to act synergistically with both sulphadiazine and pyrimethamine to restrict the growth of *T. gondii* (Table 1). We have also investigated the efficacy of glyphosate *in vivo* using a murine model, using the herbicide both alone and in combination with low doses of pyrimethamine. ND4 mice were inoculated intraperitoneally with tachyzoites of the RH strain of *T. gondii*, which under normal circumstances is lethal. The survival of mice treated with glyphosate or pyrimethamine alone was not significantly increased compared with that of untreated control mice (Fig. 1d). However, simultaneous administration of glyphosate and pyrimethamine protected about 50% of mice from death at 7 days after infection (Fig. 1d, $P < 0.05$). Mice were permitted to consume a diet that contained PABA and folate *ad libitum*. This indicates that compounds that inhibit EPSP synthase and conventional antifolates enhance the effect of each other and the *in vivo* data are consistent with the synergy demonstrated *in vitro* (Table 1). Such combinations should be useful for the treatment of toxoplasmosis. Furthermore, they could also have applications against other diseases caused by apicomplexan parasites, such as malaria. We found that glyphosate was equally effective *in vitro* against both pyrimethamine-sensitive and pyrimethamine-resistant *P. falciparum* lines (means ± s.d. of 5 replicate experiments for D6 and W2; IC_{50} 0.5 ± 0.04 and 1.1 ± 0.04 mM, IC_{90} 1.6 ± 0.1 and 2.1 ± 0.09 mM, respectively). This emphasizes the potential therapeutic value of targeting alternative enzymes of the folate biosynthetic pathway.

The discovery of the shikimate pathway in apicomplexan parasites provides new opportunities for the development of antimicrobial agents effective against these parasites. The inhibitor used in these studies, glyphosate, should be a valuable lead compound in this process. A variety of derivatives of glyphosate are currently being used to elucidate structure–function relationships for inhibitors of plant EPSP synthases¹⁸, and a similar approach could be useful for characterizing the active site of the parasite enzymes. Inhibitors of chorismate synthase¹⁹ and other enzymes within the shikimate pathway also are being developed in the search for new herbicides and antimicrobial agents effective against bacterial and fungal pathogens. These too may be useful against apicomplexan parasites. Indeed, because many other microbes that cause opportunistic infections of AIDS patients, including *Pneumocystis carinii*²⁰ and *Mycobacterium tuberculosis*²¹, also have

the shikimate pathway, there is now the exciting possibility that compounds with broad-spectrum activity could be useful against several opportunistic pathogens simultaneously infecting an individual. □

Methods

***T. gondii* in vitro.** Growth inhibition was assessed over a 4-day period as described previously²² with modifications, using human foreskin fibroblasts (HFF) infected with 10³ tachyzoites of the RH strain of *T. gondii*. For PABA competition studies, triplicate cultures were treated with the following as appropriate: glyphosate (4.5 mM), sulphadiazine (46 μM) or pyrimethamine (0.4 μM) either alone or in combination with PABA (2.5, 5, 10 or 100 μM).

Enzyme assays. The EPSP synthase assay entailed monitoring the generation of EPSP using HPLC. Reaction components were separated using a Hypersil H3APS2 HPLC column (Hichrom, Reading, UK) and a NaH_2PO_4 elution gradient (50–400 mM). Ultraviolet spectra (200–320 nm) of the column eluate were collected to identify eluants. Shikimate-3-phosphate and 5-enolpyruvyl-shikimate-3-phosphate (EPSP), synthesized enzymatically and purified to at least 95% purity as described¹², eluted after 3.9 and 6.8 min, respectively; phosphoenolpyruvate did not interfere with the EPSP detection and eluted after 5.3 min. The peaks at 215 nm were integrated; the EPSP produced was quantified using a standard curve of authentic EPSP. Parasite extracts were produced at 4 °C by suspension of pure tachyzoites in extraction buffer (50 mM Tris-HCl, pH 7.5, containing complete TM protease inhibitor cocktail (Boehringer Mannheim, 1 tablet per 50 ml buffer)), sonication 3 times for 3 s at 30-s intervals, and centrifugation at 12,000g for 15 min. The resulting supernatant was diluted sixfold with extraction buffer and loaded onto a ResourceQ column (1 ml, Pharmacia) equilibrated with extraction buffer. The bound protein was eluted in a single step using extraction buffer containing 500 mM KCl. The eluted material was used for enzyme assay. The assay mix contained 1 mM phosphoenolpyruvate, 1 mM shikimate-3-phosphate and 50 mM HEPES, pH 7.5. The reaction was started by addition of parasite extract and incubation was at 30 °C. Timed 10 μl aliquots were subjected to HPLC analysis. Protein concentrations of lysates were determined using the Lowry method.

***T. gondii* in vivo.** ND4 mice (10–12 mice per group) were infected intraperitoneally with *T. gondii* tachyzoites (RH strain). Mice were then treated with glyphosate (100 mg kg⁻¹ day⁻¹) or pyrimethamine (12.5 mg kg⁻¹ day⁻¹), alone or in combination. Treatments were administered orally in drinking water.

DNA sequencing. Clones were sequenced using a combination of gene walking and sequencing of restriction fragments and alignments of the deduced amino-acid sequences were performed^{23–27}.

***P. falciparum*.** The *in vitro* assays were conducted using a modification of the semiautomated microdilution technique for assessing antifolate antagonists²⁸. Instead of dialysed human plasma, 10% Albumax I (Gibco BRL), a serum-free substitute, was used to supplement the RPMI 1640 media. All test compounds were dissolved in DMSO and diluted 400-fold into complete media before serial dilution over 11 concentrations. Glyphosate was at 1.08 mM in serial dilutions of PABA or folic acid for the combination studies. Incubation was at 37 °C in 5% O₂, 5% CO₂ and 90% N₂ for 48 h. [³H]-Hypoxanthine incorporation was measured as described previously²⁸. Two clones of *P. falciparum* were used for the susceptibility tests²⁹. W2 is susceptible to mefloquine, but resistant to chloroquine, pyrimethamine, sulphadoxine and quinine. D6 is susceptible to chloroquine, pyrimethamine and sulphadoxine, but resistant to mefloquine.

***C. parvum*.** *C. parvum* oocysts were added at 50,000 per well to confluent MDBK F5D2 cell monolayers in 96-well microtitre plates and incubated at 37 °C under air with 8% CO₂ in medium containing glyphosate³⁰. The level of

infection after 48 h was determined by an immunofluorescence assay. *C. parvum* asexual forms were stained with rabbit anti-*C. parvum* serum and then fluorescein-conjugated goat anti-rabbit antibody. Sixteen fields were counted under $\times 10$ magnification. Data are expressed as the mean parasite count per field.

Received 26 January; accepted 20 April 1998.

1. Fichera, M. E. & Roos, D. S. A plastid organelle as a drug target in apicomplexan parasites. *Nature* **390**, 407–409 (1997).
2. McFadden, G. I., Keith, M. E., Monholland, J. M. & Lang-Unasch, N. Plastids in human parasites. *Nature* **381**, 482 (1996).
3. Kishore, G. M. & Shah, D. M. Amino acid biosynthesis inhibitors as herbicides. *Annu. Rev. Biochem.* **57**, 627–663 (1988).
4. Haslam, E. *Shikimic Acid: Metabolism and Metabolites* (John Wiley, Chichester, 1993).
5. Williamson, D. H. *et al.* The evolutionary origin of the 35kb circular DNA of *Plasmodium falciparum*: new evidence supports a possible rhodophyte ancestry. *Mol. Gen. Genet.* **243**, 249–252 (1994).
6. Kohler, S. *et al.* A plastid of probable green algal origin in apicomplexan parasites. *Science* **275**, 1485–1489 (1997).
7. Hackstein, J. H. P., Mackenstedt, U., Melhorn, H., Schubert, H. & Leunissen, J. A. M. Parasitic apicomplexans harbor a chlorophyll a-D1 complex, the potential target for therapeutic triazines. *Parasitol. Res.* **81**, 207–216 (1995).
8. Bentley, R. The shikimate pathway—a metabolic tree with many branches. *Crit. Rev. Biochem. Mol. Biol.* **25**, 307–383 (1990).
9. Grolach, J., Schmid, J. & Amrhein, N. Differential expression of tomato (*Lycopersicon esculentum* L.) genes encoding shikimate pathway isoenzymes. II. Chorismate synthase. *Plant Mol. Biol.* **23**, 707–716 (1993).
10. Haucke, V. & Schatz, G. Import of proteins into mitochondria and chloroplasts. *Cell Biol.* **7**, 103–106 (1997).
11. Steinrücken, H. C. & Amrhein, N. The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid 3-phosphate synthase. *Biochem. Biophys. Res. Commun.* **94**, 1207–1212 (1980).
12. Coggins, J. R. *et al.* The arom multifunctional enzyme (*Neurospora crassa*). *Methods Enzymol.* **142**, 325–335 (1987).
13. Ramjee, M. K., Coggins, J. R. & Thorneley, R. N. F. A continuous, anaerobic spectrophotometric assay for chorismate synthase activity that utilizes photoreduced flavin mononucleotide. *Anal. Biochem.* **220**, 137–141 (1994).
14. Deka, R., Anton, I. A., Dunbar, B. & Coggins, J. R. The characterisation of the shikimate pathway enzyme dehydroquinase from *Pisum sativum*. *FEBS Lett.* **349**, 397–402 (1994).
15. Mousdale, D. M. & Coggins, J. R. High performance liquid chromatography of shikimate pathway intermediates. *J. Chromatogr.* **329**, 268–272 (1985).
16. Duggan, P., Parker, E., Coggins, J. & Abell, C. Enzymatic synthesis of (6R) and (6S)-fluoroshikimic acids. *Bioorg. Med. Chem. Lett.* **5**, 2347–2352 (1995).
17. Dieckmann, A. & Jung, A. Mechanisms of sulfadoxine resistance in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **19**, 143–147 (1986).
18. Marzabadi, M. R. *et al.* An EPSP synthase inhibitor joining shikimate-3-phosphate with glyphosate: synthesis and ligand binding studies. *Biochemistry* **35**, 4199–4210 (1996).
19. Davies, G. M. *et al.* (6S)-6-fluoroshikimic acid, an antibacterial agent acting on the aromatic biosynthetic pathway. *Antimicrob. Agents Chemother.* **38**, 403–406 (1994).
20. Banerji, S., Lugli, E. B., Miller, R. F. & Wakefield, A. E. Analysis of genetic diversity at the arom locus in isolates of *Pneumocystis carinii*. *J. Euk. Microbiol.* **42**, 675–679 (1995).
21. Garbe, T. *et al.* The *Mycobacterium tuberculosis* shikimate pathway genes: Evolutionary relationship between biosynthetic and catabolic 3-dehydroquinases. *Mol. Gen. Genet.* **228**, 385–392 (1991).
22. Mack, D. & McLeod, R. A new micromethod to study effects of antimicrobial agents on *Toxoplasma gondii*: Comparison of sulfadoxine and sulfadiazine and study of clindamycin, metronidazole, and cyclosporin A. *Antimicrob. Agents Chemother.* **26**, 26–30 (1984).
23. White, P. J., Millar, G. & Coggins, J. R. The overexpression, purification and complete amino acid sequence of chorismate synthase from *Escherichia coli* K12 and its comparison with the enzyme from *Neurospora crassa*. *Biochem. J.* **251**, 313–322 (1988).
24. Schmidt, J., Bubunenko, M. & Subramanian, A. R. A novel operon organization involving the genes for chorismate synthase (aromatic biosynthesis pathway) and ribosomal GTPase center proteins (L11, L1, L10, L12: rplKAL) in cyanobacterium *Synechocystis* PCC 6803. *J. Biol. Chem.* **268**, 27447–27457 (1993).
25. Jones, D. G., Reusser, U. & Braus, G. H. Molecular cloning, characterization and analysis of the regulation of the *aroC* gene, encoding chorismate synthase, of *Saccharomyces cerevisiae*. *Mol. Microbiol.* **5**, 2143–2152 (1991).
26. Fleischmann, R. D. *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512 (1995).
27. Henstrand, J. M., Amrhein, N. & Schmid, J. Cloning and characterization of a heterologously expressed bifunctional chorismate synthase/flavin reductase from *Neurospora crassa*. *J. Biol. Chem.* **270**, 20447–20452 (1995).
28. Milhous, W. K., Weatherly, N. F., Bowdre, J. H. & Desjardins, R. E. *In vitro* activities and mechanisms of resistance to anti-folates and anti-malarial drugs. *Antimicrob. Agents Chemother.* **27**, 525–530 (1985).
29. Oduola, A. M. J., Weatherly, N. F., Bowdre, J. H. & Desjardins, R. E. *Plasmodium falciparum*—cloning by single erythrocyte micromanipulation and heterogeneity *in vitro*. *Exp. Parasitol.* **66**, 86–95 (1988).
30. Tzipori, S. Laboratory investigations and chemotherapy of *Cryptosporidium parvum*. *Adv. Parasitol.* **40**, 188–221 (1998).

Acknowledgements. We thank L. Mets for information about plants; D. Sibley and the Toxoplasma EST Project at Washington University for the putative chorismate synthase EST; M. Gottlieb, J. Boothroyd, K. Kim, L. Weiss, A. Lapthorn, A. Martin, M. Rust and U. Gross for discussions; and V. Aitchison and E. Hofels for assistance in preparation of this manuscript. F.R. was a recipient of a Royal Pathological Society of Great Britain and Northern Ireland Travelling Fellowship, C.W.R. was a Fulbright Scholar, and a Michael Reese Physicians' Research and Education Foundation Postdoctoral Fellow and Visiting Scholar and is a Glaxo Jack Lecturer. R.M. is the Jules and Doris Stein Research to Prevent Blindness Professor at The University of Chicago. This work was supported by the NIH NIAID TMP program, the Toxoplasmosis Research Institute, the Research to Prevent Blindness Foundation, an NIH contract from NIH NIAID (the AIDS Opportunistic Infection Drug Discovery), and the WHO-NIH-WRAIR Drug Development Programs, The Wellcome Trust and the European Union.

Correspondence and requests for materials should be addressed to R.M. (e-mail: rmcleod@midway.uchicago.edu). Genbank accession numbers for chorismate synthase sequences are U93689 and AF008549.

Regulation of actin dynamics through phosphorylation of cofilin by LIM-kinase

Silvia Arber^{*†}, Freda A. Barbayannis[‡], Hartwig Hanser^{*}, Corinna Schneider^{*}, Clement A. Stanyon[‡], Ora Bernard[‡] & Pico Caroni^{*}

^{*} Friedrich Miescher Institute, Maulbeerstrasse 66, CH-4058 Basel, Switzerland

[‡] Walter and Eliza Hall Institute of Medical Research, Victoria 3050, Australia

Cell division, cell motility and the formation and maintenance of specialized structures in differentiated cells depend directly on the regulated dynamics of the actin cytoskeleton^{1,2}. To understand the mechanisms of these basic cellular processes, the signalling pathways that link external signals to the regulation of the actin cytoskeleton need to be characterized^{2,3}. Here we identify a pathway for the regulation of cofilin, a ubiquitous actin-binding protein that is essential for effective depolymerization of actin filaments^{4–8}. LIM-kinase 1, also known as KIZ, is a protein kinase with two amino-terminal LIM motifs^{9–11} that induces stabilization of F-actin structures in transfected cells. Dominant-negative LIM-kinase1 inhibits the accumulation of the F-actin. Phosphorylation experiments *in vivo* and *in vitro* provide evidence that cofilin is a physiological substrate of LIM-kinase 1. Phosphorylation by LIM-kinase 1 inactivates cofilin, leading to accumulation of actin filaments. Constitutively active Rac augmented cofilin phosphorylation and LIM-kinase 1 autophosphorylation whereas phorbol ester inhibited these processes. Our results define a mechanism for the regulation of cofilin and hence of actin dynamics *in vivo*. By modulating the stability of actin cytoskeletal structures, this pathway should play a central role in regulating cell motility and morphogenesis.

The closely related LIM-kinase proteins 1 (refs 9, 10) and 2 (ref. 12) (LIMK-1 and LIMK-2) are serine kinases for which no physiological substrates have been identified. One potential clue to their function is that LIMKs have two LIM domains¹¹ at their amino-terminal end; these domains are protein-binding motifs that are frequently found in cytosolic proteins which interact with the actin cytoskeleton^{11,13,14}. When analysing the actin cytoskeleton of cells overexpressing LIMK-1, we noticed that F-actin structures visualized with rhodamine-conjugated phalloidin were markedly more prominent (Fig. 1a) and less sensitive to depolymerization in the presence of cytochalasin B (Fig. 1b) than they were in wild-type cells. In similar experiments we observed no alterations in the arrangement of microtubules or intermediate filaments (data not shown), and cells transfected with a green fluorescent protein control exhibited no alterations in the actin cytoskeleton (Fig. 1d).

As shown in Fig. 1c, d, expression of LIMK-1 or LIMK-2 constructs lacking the LIM domains (these constructs are called K1 or K2) induced dramatic accumulation of the actin cytoskeleton into large clumps, with clearance of any residual F-actin signal. LIMK-1 has a splice variant, LIMK1-short (LIMK1(-)), that lacks 20 amino acids in the catalytic domain⁹ and has no kinase activity (Fig. 2b). When compared to non-transfected cells, cells overexpressing LIMK1-short or K1-short exhibited distinctly reduced phalloidin staining (Fig. 1c–e), indicating that the effects of LIMK-1 and K1 on the actin cytoskeleton depend on the kinase activity of LIMK-1, and suggesting that LIMK1-short and K1-short may reduce F-actin contents by inhibiting endogenous LIM-kinase activity. In cotransfection experiments, both LIMK1-short and K1-short effectively inhibited the effects of LIMK1 (Fig. 1d), indicating that

[†] Present address: HHMI, Department of Biochemistry and Molecular Biophysics, Hammer Health Sciences, Columbia University, New York, New York 10032, USA.

LIMK1-short and K1-short can act as dominant-negative inhibitors of LIMK-1.

To monitor the effects of LIMK-1 on the formation of functionally defined actin structures, we expressed this protein in a neuronal cell line. The leading edge of neuronal growth cones accumulates characteristic F-actin-based structures such as lamellipodia and filopodia that mediate growth-cone motility and guidance¹⁵. As shown in Fig. 1e, overexpression of LIMK-1 greatly enhanced the contents of such structures in PC12 growth cones. In contrast, overexpression of LIMK1-short substantially reduced F-actin accumulation (Fig. 1e), supporting the idea that endogenous LIMK-kinase activity modulates the accumulation of actin-based structures in many cell types.

The dramatic and highly characteristic effects of LIMK-1 on the actin cytoskeleton indicated that this protein may be specifically involved in cytoskeletal regulation, possibly by reducing depolymerization of actin filaments. As cofilin hypomorphs in

*Drosophila*¹⁶ and cofilin mutants in yeast⁵ display F-actin accumulation and clumping reminiscent of the effects induced by LIMKs, we reasoned that LIMKs may act by promoting the phosphorylation and thus inactivation¹⁷⁻¹⁹ of cofilin. To determine whether LIMK-1 promotes phosphorylation of cofilin, we carried out phosphorylation experiments *in situ*. As shown in Fig. 2a, LIMK-1 promoted substantial phosphorylation of cofilin, but not of the Ser 3 → Ala cofilin mutant (cofilin (S3A)) that lacked the physiological

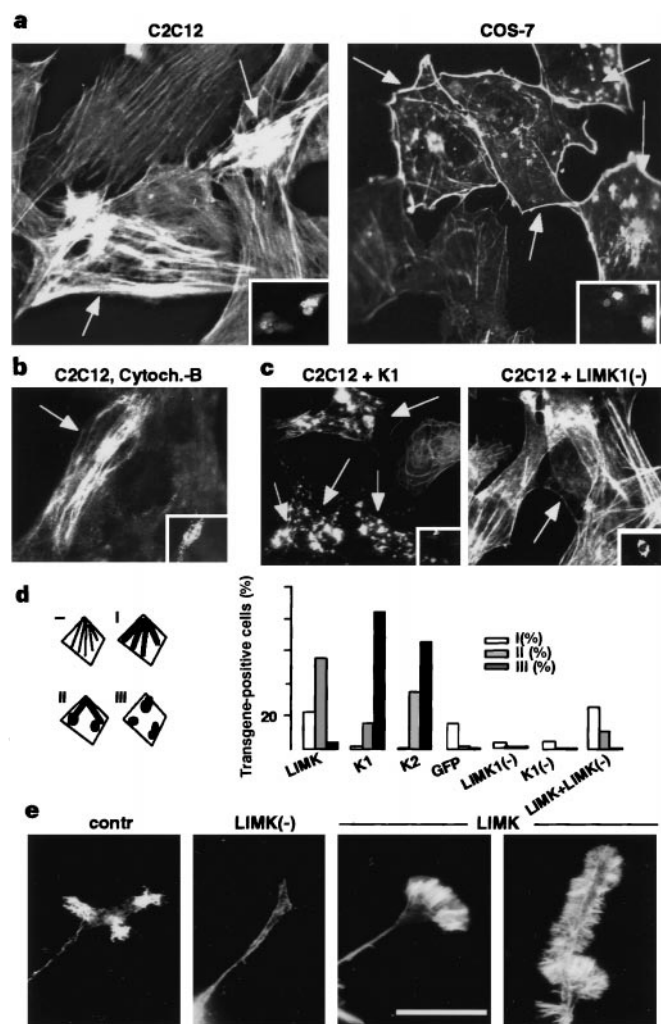


Figure 1 LIMK-1 induces accumulation of the actin cytoskeleton in transfected cells. Cells were double stained for the Myc epitope (insets) and for F-actin (RITC-phalloidin staining). Transgene-expressing cells stain for the Myc epitope and are indicated by arrows. **a**, Actin cytoskeleton accumulation in LIMK1-overexpressing C2C12 and COS-7 cells. **b**, Protection against cytochalasin-B-induced depolymerization in LIMK1-overexpressing cells. **c**, Identification of constitutively active (K1) and dominant-negative (LIMK1(-)) forms of LIMK-1. **d**, Quantitative analysis of C2C12 data as shown in **a** and **c**; F-actin categories (I, II and III) are shown on the left (see Methods). GFP, green fluorescent protein control. **e**, LIMK-1 and LIMK1-short (LIMK1(-)) modulate the accumulation of actin structures in PC12 growth cones (phalloidin stainings). Bar shown in **e** represents: **a**, left, 17 μ m; **a**, right, 26 μ m; **b**, 23 μ m; **c**, left, 35 μ m; **c**, right, 24 μ m; **e**, 11.5 μ m.

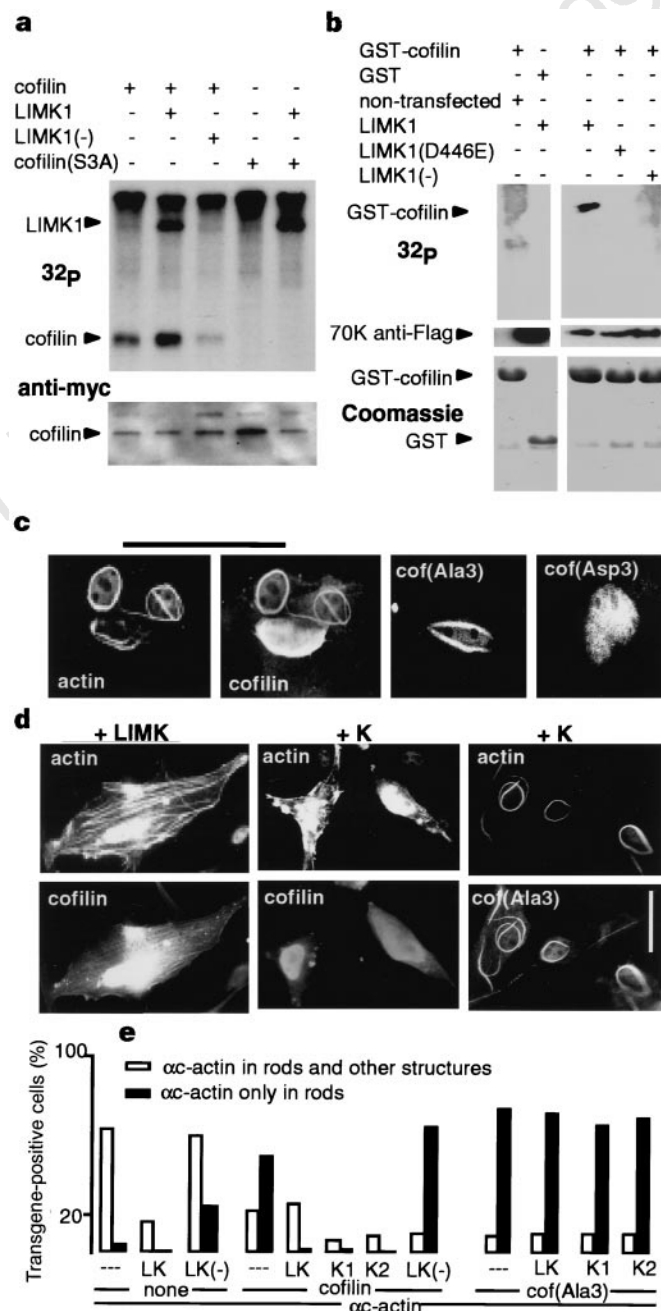


Figure 2 LIMK-1 promotes phosphorylation of cofilin *in vivo* and *in vitro*, and inactivation of cofilin *in situ*. **a**, Phosphorylation of cofilin in U2-SO cells. **b**, *In vitro* phosphorylation of recombinant GST-cofilin by LIMK1. **c**, Actin-cofilin rod formation in the presence of α -cardiac actin. Double transfections of C2C12 cells with α -cardiac actin and either wild-type cofilin or its phosphorylation-site mutants (staining for cofilin mutant shown). Cof(Ala3), cofilin (S3A); cof(Asp3), cofilin (S3D). **d**, LIMK-1 and K1 suppress rod formation in the presence of wild-type, but not non-phosphorylatable, cofilin (S3A). Experimental conditions as in **c**. **e**, Quantitative analysis of data shown in **c**, **d**. Bar shown in **d** represents: **c**, 26 μ m; **d**, LIMK1, 17 μ m; **d**, K, 28 μ m. LIMK1(-), LIMK1-short; α -actin, α -cardiac actin.

phosphorylation site¹⁹. Consistent with the results shown in Fig. 1c–e, overexpression of LIMK1-short reduced cofilin phosphorylation (Fig. 2a), indicating that this inactive form of LIMK-1 may inhibit an endogenous cofilin kinase activity. To obtain more direct evidence that cofilin is a direct substrate for LIMK-1, we carried out *in vitro* phosphorylation experiments with recombinant glutathione S-transferase (GST)–cofilin. As shown in Fig. 2b, a cell extract immunopurified from cells overexpressing LIMK-1 phosphorylated cofilin. Corresponding fractions containing Flag–LIMK1-short or Flag–LIMK-1 (D446G), which contains a mutation of Asp 446 to Glu and which lacks kinase activity, did not phosphorylate cofilin. These results support the idea that cofilin is a physiological substrate of LIMK-1.

Does LIMK-1 effectively inactivate cofilin in living cells? To address this issue, we developed an assay based on the formation of prominent actin–cofilin rod^{4,20} structures in cells overexpressing α -cardiac actin and active cofilin (Fig. 2c; see Methods). A cofilin (S3A) mutant that cannot be inactivated by phosphorylation had similar effects to those of cofilin, whereas a cofilin (S3D) mutant that mimics phosphorylated cofilin failed to accumulate at the rods (Fig. 2c). Transfection of wild-type cofilin together with LIMK-1 or K1, but not with LIMK1-short (nor K1-short; data not shown), inhibited actin–cofilin rod formation (Fig. 2d, e). Moreover, the inhibitory effects of LIMK-1 or kinase-only constructs on rod formation were completely suppressed when the non-phosphorylatable mutant cofilin (S3A) was transfected instead of wild-type cofilin (Fig. 2d, e). These results are biological evidence that LIMK-1 induces substantial inactivation of cofilin through phosphorylation of cofilin at S3.

To determine whether the effects of LIMK-1 on the actin cytoskeleton are mediated by the inactivation of cofilin, we carried out co-transfection experiments with wild-type cofilin and cofilin (S3A). Overexpression of the cofilin constructs induced pronounced loss of actin filament structures in transfected cells (Fig. 3, left panels). Transfection of LIMK-1 together with cofilin produced an intermediate result, with most cells exhibiting substantially less F-actin accumulation than in the presence of LIMK alone (Fig. 3, upper row). This result is consistent with the idea that

LIMK-1 inhibits the effects of cofilin on the actin cytoskeleton by reducing the pool of active cofilin. The effects of the much more potent kinase-only constructs K1 or K2 were affected very little by increased levels of cofilin expression (Fig. 3). In contrast, co-transfection with cofilin (S3A) completely abolished any detectable effect of LIMK-1 or of the kinase-only proteins on the actin cytoskeleton (Fig. 3, lower row). These results, together with those of the phosphorylation and rod-formation experiments, indicate that LIMK-1 and the kinase-only proteins induce accumulation and stabilization of actin filaments through phosphorylation and inhibition of cofilin function.

To identify signalling pathways that may be involved in the regulation of cofilin phosphorylation by LIM-kinases, we investigated experimental conditions that lead to actin cytoskeleton rearrangements. The addition of the phorbol ester phorbol 12-myristoyl-13-acetate (PMA) to 3T3 fibroblasts induced a dramatic

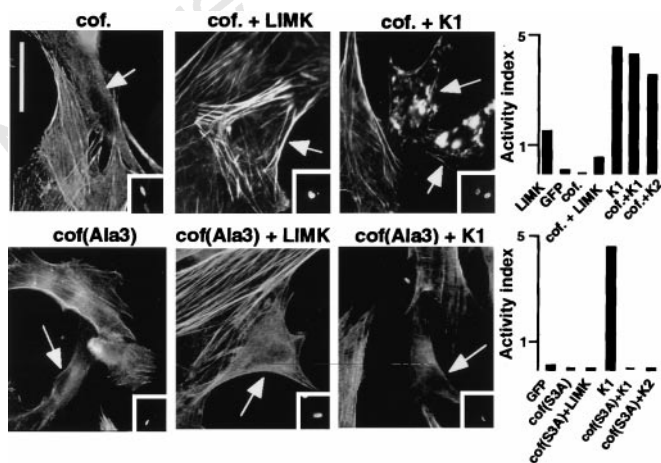


Figure 3 LIM-kinases stabilize the actin cytoskeleton through inactivation of cofilin. Transfection and staining of C2C12 cells are as described in Fig. 1. Top row, transfections with Myc-tagged wild-type cofilin (cof.), with or without LIMK-1 or K1 (right, quantitative analysis of similar experiments. Activity index N = percentage of cells with F-actin structures in category I + 2 (percentage of category II) + 5 (percentage of category III); $0 < N < 5$; see Methods for F-actin categories). Bottom row, corresponding experiments with Myc-tagged cofilin (S3A). The non-phosphorylatable mutant cofilin (S3A) completely suppressed any effect of LIMK-1 or of the activated kinase-only proteins on the actin cytoskeleton. Bar, 25 μ m.

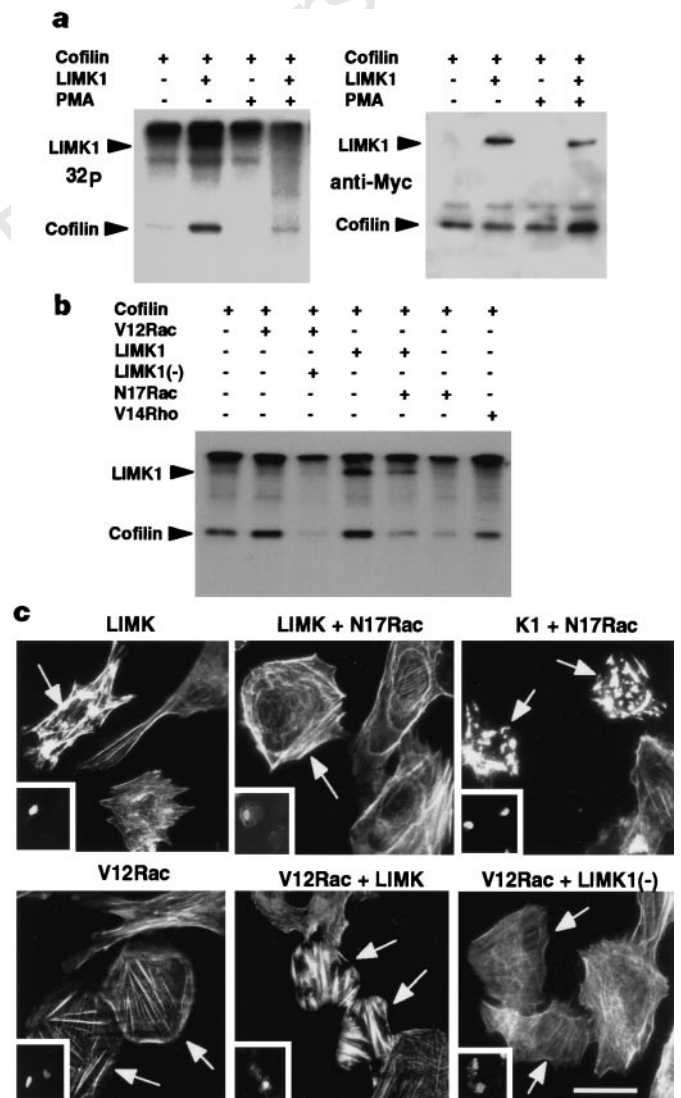


Figure 4 PMA and activated Rac affect LIMK-1 activity and cofilin phosphorylation. **a**, Phorbol ester (15 min) decreases the phosphorylation of LIMK-1 and cofilin in transfected 3T3 cells. **b**, Elevated phosphorylation of cofilin in U2-50 cells in the presence of Rac V12. Parallel immunoblots showed comparable expression levels of LIMK-1 ($\pm$ Rac N17) and of Rac V12 ($\pm$ LIMK1(-), that is, LIMK1-short) (data not shown). **c**, Dominant-negative Rac N17 suppresses the effects of LIMK-1, but not those of kinase only K1; the effects of Rac V12 are potentiated by overexpression of LIMK-1, and are inhibited by dominant-negative LIMK1(-). Phalloidin and anti-Myc (insets) labellings are as described in Fig. 1. Bar, 13 μ m.

loss of stress fibres and the formation of membrane ruffles. These changes were complete within 3–6 min in non-transfected cells, but took up to 10–15 min in cells overexpressing LIMK-1 (data not shown). In parallel experiments, PMA induced a marked reduction in both cofilin phosphorylation and LIMK-1 autophosphorylation (Fig. 4a; 15 min), indicating that, under these experimental conditions, signals initiated by protein kinase C lead to inhibition of LIMK-1 activity and dephosphorylation of cofilin.

We also carried out co-transfections with constitutively active and dominant-negative constructs containing actin-cytoskeleton regulators from the Rho family of small G proteins³. Constitutively active Rac V12 (but not Rho V14) enhanced the phosphorylation of cofilin in transfected cells, whereas dominant-negative Rac N17 (but not Rho N19, nor Cdc42 N17) reduced it (Fig. 4b). Expression levels of the small-G-protein constructs were similar. These results indicate that Rac V12 is involved in the endogenous network that controls cofilin phosphorylation levels. Rac N17 reduced the phosphorylation of cofilin in the presence of LIMK-1, and LIMK1-short reduced the effect of Rac V12 on cofilin phosphorylation (Fig. 4b), indicating that Rac V12 and LIMK-1 are on interrelated pathways that increase the proportion of phosphorylated inactive cofilin. Consistent with this interpretation, Rac N17 suppressed the effects of LIMK-1, but not of K1, on F-actin accumulation (Fig. 4c). C2C12 cells overexpressing Rac V12 had a characteristic round shape, with pronounced lamellipodia and parallel arrays of well-defined stress fibres (Fig. 4c, bottom left). Transfection of Rac V12 with LIMK-1 resulted in a dramatic accentuation and contraction of the lamellar and stress-fibre F-actin system (Fig. 4c, bottom centre), whereas dominant-negative LIMK1-short greatly reduced the prominence of F-actin structures, indicating that the effects of Rac V12 on the actin cytoskeleton are accentuated by endogenous LIM-kinase activity.

By providing evidence that one effector pathway of Rac leads to elevated phosphorylation and inactivation of cofilin, thus increasing the effects of LIM-kinases, these findings link the regulatory activity of LIMK-1 to the effects of Rac, and hence to major signalling pathways for actin cytoskeleton regulation. One possibility is that Rac V12 lies upstream of LIMK-1 in a signalling pathway promoting cofilin phosphorylation. The fact that the kinase-only construct K1 was apparently not affected by Rac N17 might reflect loss of Rac-dependent regulation in the absence of the LIM motifs.

Several observations indicate that the pool of active cofilin is vital for regulation of the actin cytoskeleton and is controlled by phosphorylation^{5,16,18,21}. We propose that local concentrations of active cofilin are controlled by cofilin kinases (that is, LIM-kinases) and phosphatases, and that antagonistic signalling pathways may converge on cofilin phosphorylation and dephosphorylation reactions to control pools of active cofilin in the cell, and thus actin filament dynamics. Consistent with the fact that such regulation is crucial for actin-based motility, fibroblast migration in a monolayer-wounding assay *in vitro* was suppressed by overexpression of either LIMK-1 or cofilin (data not shown). Expression of both LIMK-1 and cofilin partly rescued cell motility, supporting the idea that actin-based motility may involve local modulation of cofilin-phosphorylation pathways. LIMK-1 is expressed at comparatively high levels in neurons and their processes, where it may also be regulated by the inactive kinase splice form LIMK1-short⁹. We found that LIMK-1 and LIMK1-short can form homodimers and heterodimers in transfected COS cells (data not shown), suggesting a mechanism by which LIMK1-short may negatively regulate the activity of LIMK-1. In the nervous system, regulation of LIMK-1 activity may affect the formation and/or stability of essential actin-based structures in growth cones and postsynaptic densities.

Methods

Reagents, cell culture and transfections. Cell lines were cultured and transfected as described¹³. PC12Y cells grown on acid-washed glass coverslips

were transfected, and then maintained for 2 days in the presence of 4% FCS and 50 ng ml⁻¹ nerve growth factor. Where added, cytochalasin B (Sigma) was 50 μ M and PMA (Sigma) was 10 nM. Most experiments were analysed¹³ 16–18 h after transfection, that is, 3–5 h after first detection of transgene expression. Antibodies were against Myc (monoclonal antibody MYCL-9E10 from American Type Cell Culture Collection, or rabbit antiserum against synthetic Myc peptide), GLVVMNIT¹³, His₆ (monoclonal antibody, Qiagen) or VSV²². As most of the effects described here were easily apparent, double staining of particular combinations of transgenes (for example α -cardiac actin and cofilin, Fig. 2d, e) was routinely sufficient for the analysis.

To ensure that the effects observed were due to expression of all transgenes in a transfected cell, the analysis also included: first, for triple transfections, double staining for different pairs of constructs in parallel transfections; and second, triple staining with rhodamine isothiocyanate (RITC)-phalloidin and secondary antibodies labelled with lucifer yellow and cascade blue (Molecular Probes). Relative expression levels varied significantly, but <10% of the cells expressed no detectable levels of one particular construct. To quantify patterns of labelling of the actin cytoskeleton, all transgene-positive cells from randomly selected fields ($\times 500$ magnification) were scored, for a total of 100 cells each. The categories used for the analysis (Fig. 1d, left) were: –, no differences in F-actin labelling between transfected and non-expressing cells; I, brighter fibres but no clumps; II, bright fibres and clumps; III, only clumps (see examples in Fig. 1d).

3T3 fibroblast monolayers were transfected and 16 h later linear wounds were made in the cultures with a plastic pipette tip. Three hours later, when cells were processed for immunocytochemistry, wound outlines were still clearly detectable, and cell densities in the wound areas were 10–20%.

Phosphorylation experiments. Eighteen hours after transfection, U2-SO cultures were incubated for 6 h in phosphate-free medium, with 0.2 mCi ml⁻¹ of [³²P]orthophosphate. Cells were then lysed for 5 min at 4 °C in isotonic NaCl/Na phosphate buffer pH7.2, with 1% NP40 and phosphatase inhibitors. Solubilized proteins were incubated with anti-Myc-protein-G agarose (Sigma) and washed three times with lysis buffer. Bound proteins were eluted by boiling in SDS-PAGE sample buffer, and comparable amounts of cell equivalent were added to each lane. In parallel experiments, equivalent samples were processed for immunoblots. For *in vitro* phosphorylation experiments, Flag-tagged LIMK-1 constructs were transfected into COS cells and immunopurified with anti-Flag M2 affinity beads (Eastman Kodak). After five washes, the beads were resuspended in 20 μ l kinase buffer (20 mM HEPES pH 7.4, 150 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂, 50 mM NaF, 1 mM Na₃VO₄, 10 mM NaP₂O₇), with [γ -³²P]ATP and 5 μ g purified bacterially expressed GST-cofilin (a gift from D. Watters) or GST, and incubated at 37 °C for 30 min.

Complementary DNA constructs. Constructs were as follows: mouse kiz1/LIMK-1 and kiz1s/LIMK-1s (ref. 9); spacer-K1 (LIMK1_{124–632}); K1 (LIMK1_{393–632}); K2 (amino acid 393 to stop codon from rat LIMK-2); rat cofilin; vesicular stomatitis virus (VSV)/ α -cardiac actin²²; human Rac V12, Rho V14, Rac N17, Rho N19 and Cdc42 N17 (gifts from K. Ballmer and A. Hall); and GFP (pEGFP, Clontech). The Rac, Rho and Cdc42 constructs carried an N-terminal Myc tag (amino-acid sequence EQKLISEEDLN). LIM-kinase cDNA constructs carried an N-terminal Myc tag and carboxy-terminal GLVVMNIT tag¹³, a C-terminal GLVVMNIT tag, a C-terminal Myc tag, an N-terminal Flag tag, or a C-terminal His₆ tag; cofilin constructs carried either a Myc tag (N- or C-terminal) or a C-terminal His₆ tag. Mutations in the cofilin constructs were verified by sequencing.

Received 24 February; accepted 15 April 1998.

1. Michison, T. J. & Cramer, L. P. Actin-based cell motility and cell locomotion. *Cell* **84**, 371–379 (1996).
2. Welch, M. D., Mallavarapu, A., Rosenblatt, J. & Mitchison, T. J. Actin dynamics *in vivo*. *Curr. Opin. Cell Biol.* **9**, 54–61 (1997).
3. Tapon, N. & Hall, A. Rho, Rac and cdc42 GTPases regulate the organization of the actin cytoskeleton. *Curr. Opin. Cell Biol.* **9**, 86–92 (1997).
4. Moon, A. & Drubin, D. G. The ADF/cofilin proteins: stimulus-responsive modulators of actin dynamics. *Mol. Biol. Cell* **6**, 1423–1431 (1995).
5. Lappalainen, P. & Drubin, D. G. Cofilin promotes rapid actin filament turnover *in vivo*. *Nature* **388**, 78–82 (1997).
6. Rosenblatt, J., Agnew, B. J., Abe, H., Bamburg, J. R. & Mitchison, T. J. *Xenopus* actin depolymerizing factor/cofilin (XAC) is responsible for the turnover of actin filaments in *Listeria monocytogenes* tails. *J. Cell Biol.* **136**, 1323–1332 (1997).
7. Carlier, M.-F. *et al.* Actin depolymerizing factor (ADF/cofilin) enhances the rate of filament turnover: implication in actin-based motility. *J. Cell Biol.* **136**, 1307–1323 (1997).
8. Theriot, J. A. Accelerating on a treadmill: ADF/cofilin promotes rapid actin filament turnover in the dynamic cytoskeleton. *J. Cell Biol.* **136**, 1165–1168 (1997).

9. Bernard, O., Ganiatsas, S., Kannourakis, G. & Dringen, R. Kiz-1, a protein with LIM zinc finger and kinase domains, is expressed mainly in neurons. *Cell Growth Differ.* **5**, 1159–1171 (1994).
10. Mizuno, K. *et al.* Identification of a human cDNA encoding a novel protein kinase with two repeats of the LIM/double zinc finger motif. *Oncogene* **9**, 1605–1612 (1994).
11. David, I. B., Toyama, R. & Taira, M. LIM domain proteins. *C.R. Acad. Sci.* **318**, 295–306 (1995).
12. Okano, I. *et al.* Identification and characterization of a novel family of serine/threonine kinases containing two N-terminal LIM motifs. *J. Biol. Chem.* **270**, 31321–31330 (1995).
13. Arber, S. & Caroni, P. Specificity of single LIM motifs in targeting and LIM/LIM interactions *in situ*. *Genes Dev.* **10**, 289–300 (1996).
14. Arber, S. *et al.* MLP-deficient mice exhibit a disruption of cardiac cytoarchitectural organization, dilated cardiomyopathy, and heart failure. *Cell* **88**, 393–403 (1997).
15. Tanaka, E. & Sabry, J. Making the connection: cytoskeletal rearrangements during growth cone guidance. *Cell* **83**, 171–176 (1995).
16. Gunsalus, K. C. *et al.* Mutations in twinstar, a *Drosophila* gene encoding a cofilin/ADF homologue, result in defects in centrosome migration and cytokinesis. *J. Cell Biol.* **131**, 1243–1259 (1995).
17. Morgan, T. E., Lockerbie, R. O., Minamide, L. S., Browning, M. D. & Bamberg, J. R. Isolation and characterization of a regulated form of actin depolymerizing factor. *J. Cell Biol.* **122**, 623–633 (1993).
18. Abe, H., Obinata, T., Minamide, L. S. & Bamberg, J. R. *Xenopus laevis* actin-depolymerizing factor/cofilin: a phosphorylation-regulated protein essential for development. *J. Cell Biol.* **132**, 871–885 (1996).
19. Agnew, B. J., Minamide, L. S. & Bamberg, J. R. Reactivation of phosphorylated actin depolymerizing factor and identification of the regulatory site. *J. Biol. Chem.* **270**, 17582–17587 (1995).
20. Nagoaka, R., Abe, H., Kusano, K. & Obinata, T. Effects of cofilin on actin filamentous structures in cultured muscle cells. *J. Cell Sci.* **108**, 581–593 (1995).
21. Nebel, G., Meurer, S. C. & Samstag, Y. Dephosphorylation of serine 3 regulates nuclear translocation of cofilin. *J. Biol. Chem.* **271**, 26276–26280 (1996).
22. vonArx, P., Bantle, S., Soldati, T. & Perriard, J.-C. Dominant negative effect of cytoplasmic actin isoproteins on cardiomyocyte cytoarchitecture and function. *J. Cell Biol.* **131**, 1759–1773 (1995).

Acknowledgements. We thank J. Adams, T. Jessell, W. Kerk, U. Müller and G. Thomas for comments on the manuscript. S.A. and H.H. were supported by grants from the Swiss Foundation for Research on Muscle Diseases. O.B. was supported by the Australian NH-MRC, and the Australian Cooperative Research Centre Scheme, and O.B. and C.A.S. were supported by the National Multiple Sclerosis Society of Australia.

Correspondence and requests for materials should be addressed to P.C. (e-mail: caroni@fmi.ch).

Cofilin phosphorylation by LIM-kinase 1 and its role in Rac-mediated actin reorganization

Neng Yang^{*†}, Osamu Higuchi^{*†}, Kazumasa Ohashi^{*†},
Kyoko Nagata^{*}, Atsushi Wada[‡], Kenji Kangawa[§],
Eisuke Nishida[‡] & Kensaku Mizuno^{*¶}

^{*} Department of Biology, Faculty of Science, Kyushu University, Fukuoka 812-8581, Japan

[‡] Department of Biophysics, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

[§] National Cardiovascular Center Research Institute, Suita, Osaka 565-8565, Japan

[¶] Inheritance and Variation, PRESTO, Japan Science and Technology Corporation, Kyoto 619-0237, Japan

[†] These authors contributed equally to this work.

Rac is a small GTPase of the Rho family that mediates stimulus-induced actin cytoskeletal reorganization to generate lamellipodia^{1–5}. Little is known about the signalling pathways that link Rac activation to changes in actin filament dynamics. Cofilin is known to be a potent regulator of actin filament dynamics^{6–10}, and its ability to bind and depolymerize actin is abolished by phosphorylation of serine residue at 3 (refs 11, 12); however, the kinases responsible for this phosphorylation have not been identified. Here we show that LIM-kinase 1 (LIMK-1), a serine/threonine kinase containing LIM and PDZ domains^{13–16}, phosphorylates cofilin at Ser3, both *in vitro* and *in vivo*. When expressed in cultured cells, LIMK-1 induces actin reorganization and reverses cofilin-induced actin depolymerization. Expression of an inactive form of LIMK-1 suppresses lamellipodium formation induced by Rac or insulin. Furthermore, insulin and an active form of Rac increase the activity of LIMK-1. Taken together, our results indicate that LIMK-1 participates in Rac-mediated actin cytoskeletal reorganization, probably by phosphorylating cofilin.

To examine the cellular functions of LIMK-1, we searched for proteins that associate with it. We loaded extracts of rat brain onto a column that was coupled with glutathione *S*-transferase (GST) or a GST–LIMK1 fusion protein, and analysed bound proteins. We detected a major protein of relative molecular mass 42,000 (*M*_r 42K) (p42) in eluates from a GST–LIMK1 column but not from a control GST column (Fig. 1a). Sequences of eight peptides derived from p42 proved to be identical to sequences from rat β -actin (data not shown). Immunoblot analysis further confirmed that the LIMK1-binding protein p42 is rat β -actin (Fig. 1b). To determine whether LIMK-1 would bind directly to filamentous (F)-actin, we carried out an *in vitro* co-sedimentation analysis. When GST–LIMK1 or GST was centrifuged at high speed in the presence or absence of F-actin, GST–LIMK1 was pelleted only in the presence of F-actin, whereas control GST was not pelleted even in the presence of F-actin (Fig. 1d). These results indicate that LIMK-1 directly binds to F-actin. To determine the domain through which LIMK-1 binds to F-actin, we sedimented GST fusion proteins containing a LIM domain (GST–LIM), a PDZ domain (GST–PDZ) or a protein

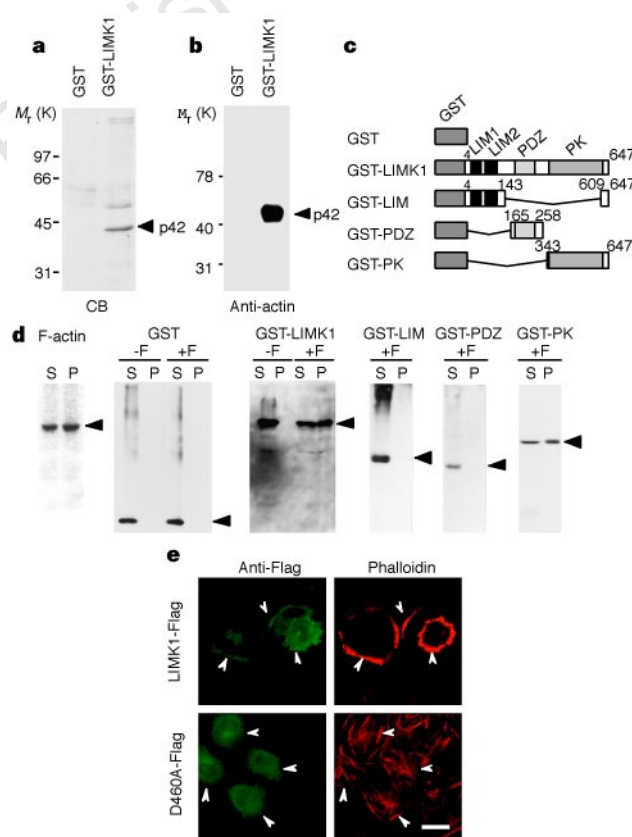


Figure 1 LIMK-1 binds to F-actin and induces actin reorganization. **a**, Coomassie blue staining of proteins bound to GST or GST–LIMK1. Arrowhead indicates a 42K protein (p42) that is bound specifically to GST–LIMK1. **b**, Immunoblot analysis, using an anti- β -actin antibody, of proteins bound to GST or GST–LIMK1. **c**, GST fusion proteins. Amino-acid residue numbers are indicated at the top of the bars. PK, protein kinase. **d**, F-actin co-sedimentation assay. GST fusion proteins were incubated with or without F-actin (+F or –F). After centrifugation, proteins recovered in the pellets (P) and the supernatants (S) were analysed by immunoblotting with anti-GST (GST, GST–PDZ) or anti-LIMK1 (GST–LIMK1, GST–LIM, GST–PK) antibodies. As a control, F-actin alone was centrifuged and recovered actin was stained with Coomassie blue. Arrowheads indicate the positions of actin or GST fusion proteins. **e**, Actin reorganization induced by LIMK-1. HeLa cells transfected with plasmids for Flag-tagged LIMK-1 or LIMK-1(D460A) were stained with anti-Flag antibody (left) and rhodamine–phalloidin for F-actin (right). Arrowheads indicate cells that express LIMK-1 or LIMK-1(D460A), as assigned by anti-Flag immunostaining. Scale bar, 20 μ m.

kinase domain (GST-PK) (Fig. 1c) with F-actin. GST-PK precipitated with F-actin, but GST-LIM and GST-PDZ did not (Fig. 1d). We concluded that LIMK-1 associates with F-actin through its carboxy-terminal region, which contains a protein kinase domain.

Direct interaction of LIMK-1 with F-actin indicates that LIMK-1 may regulate actin cytoskeletal organization. We therefore transfected LIMK-1 complementary DNA into HeLa cells and visualized actin organization by rhodamine-phalloidin staining. Marked alterations in actin organization, namely the accumulation of polymerized actin at the cell periphery, occurred in LIMK1-expressing cells compared with surrounding cells that did not express LIMK-1 (Fig. 1e). LIMK-1 was enriched in regions of the polymerized actin. In contrast, expression of a kinase-inactive mutant of LIMK-1, LIMK-1(D460A), in which the catalytic residue Asp 460 is replaced with Ala, had no apparent effect on actin organization (Fig. 1e). Similar results were obtained when LIMK-1 or LIMK-1(D460A) was expressed in COS cells. Thus, LIMK-1 can induce actin reorganization, the function of which depends on its catalytic kinase activity.

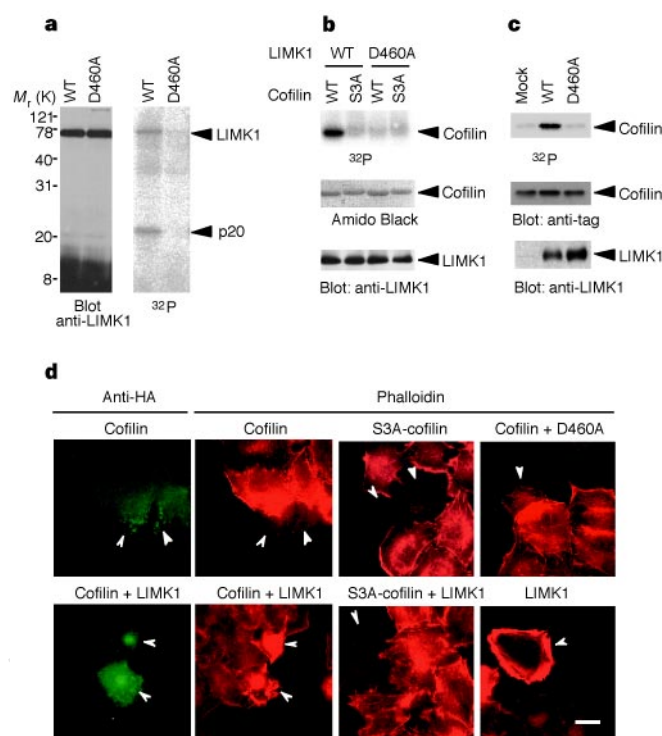


Figure 2 Phosphorylation of cofilin by LIMK-1. **a**, LIMK-1 phosphorylates a 20K protein. Lysates of COS cells transfected with plasmids encoding wild-type LIMK-1 (WT) or LIMK-1(D460A) were immunoprecipitated with anti-LIMK1 antibody. Bound proteins were eluted with antigenic peptide, reacted with kinase *in vitro*, and analysed by immunoblotting (left) or autoradiography (right). **b**, *In vitro* phosphorylation of cofilin. (His)₆-tagged cofilin (WT) or S3A-cofilin was incubated with [γ-³²P]ATP and LIMK-1 or LIMK-1(D460A) prepared from COS cells, run on SDS-PAGE, and analysed by autoradiography, amido-black staining or immunoblotting with anti-LIMK1 antibody. **c**, *In vivo* phosphorylation of cofilin. COS cells were transfected with plasmids encoding Sky-tagged cofilin alone (mock) or Sky-tagged cofilin with either LIMK-1 (WT) or LIMK-1(D460A). These cells were labelled with [³²P]phosphoric acid. Cell lysates were immunoprecipitated with anti-Sky-tag or anti-LIMK1 antibody, and analysed by autoradiography and immunoblotting. **d**, LIMK-1 inhibits actin depolymerization induced by cofilin, but not that induced by S3A-cofilin. HeLa cells injected with plasmid encoding haemagglutinin (HA)-tagged cofilin or S3A-cofilin, with or without plasmid for Flag-tagged LIMK-1 or LIMK-1(D460A), were stained with anti-HA (or anti-Flag) antibody and rhodamine-phalloidin. Arrowheads indicate the injected cells, as assigned by anti-HA or anti-Flag immunostaining. Scale bar, 20 μm.

To elucidate the mechanism by which this kinase leads to changes in actin organization, we searched for the substrate of LIMK-1 in LIMK-1 immunoprecipitates obtained from lysates of COS cells transfected with LIMK-1 cDNA. An *in vitro* kinase reaction with [γ-³²P]ATP revealed two major phosphorylated bands with *M_r* values of 70K and 20K (Fig. 2a). These ³²P-labelled bands were not detected in immunoprecipitates of a kinase-inactive LIMK-1(D460A), indicating that these proteins were phosphorylated by LIMK-1. Immunoblot analysis showed that the 70K band represented autophosphorylated LIMK-1. We proposed that the 20K phosphorylated protein might be cofilin, as cofilin is a 20K protein with actin-binding and -depolymerization activities⁶⁻⁸ that are abolished by phosphorylation at Ser^{31,12}. To test this hypothesis, we used recombinant (His)₆-tagged cofilin and S3A-cofilin, in which Ser3 is replaced with alanine, as substrates in an *in vitro* kinase reaction. LIMK-1 efficiently phosphorylated cofilin, but not S3A-cofilin (Fig. 2b). LIMK-1(D460A) did not phosphorylate either protein. ³²P-phosphorylated cofilin had no F-actin-binding ability (data not shown). These findings indicate that LIMK-1 may phosphorylate cofilin specifically at Ser 3 and that this phosphorylation inactivates cofilin.

To determine whether LIMK-1 can induce phosphorylation of cofilin *in vivo*, we expressed epitope-tagged cofilin in COS cells alone or with LIMK-1 or LIMK-1(D460A). We then labelled the cells with [³²P]phosphoric acid. Expression of cofilin with LIMK-1, but not with LIMK-1(D460A), increased the level of ³²P incorporation into cofilin (Fig. 2c), indicating that LIMK-1 can phosphorylate cofilin in cultured cells. When we injected cDNAs encoding cofilin and S3A-cofilin into HeLa cells, actin depolymerized, as measured by rhodamine-phalloidin staining (Fig. 2d). When LIMK-1 was injected with cofilin, cofilin-induced actin depolymerization was repressed, and polymerized actin structures were observed. In contrast, actin depolymerization induced by S3A-cofilin (a non-phosphorylatable form of cofilin) was not affected by injection of S3A-cofilin with LIMK-1 (Fig. 2d). A kinase-inactive LIMK-1(D460A) did not reverse cofilin-induced actin depolymerization. These results are consistent with the idea that LIMK-1 can phosphorylate cofilin at Ser 3 and inhibit the actin-depolymerizing activity of cofilin *in vivo*.

The Rho family of small GTPases, including Rho, Rac and Cdc42, is known to regulate actin cytoskeletal reorganization¹. As seen in fibroblasts^{3,4,17}, microinjection of plasmids encoding activated forms of Rho (RhoV14), Rac (RacV12) and Cdc42 (Cdc42V12) into HeLa cells induced distinct patterns of actin reorganization, such as stress fibre formation by RhoV14 and lamellipodium formation by RacV12 (Fig. 3a, upper panels). As LIMK-1 induces changes in actin organization, this kinase may be involved in signalling pathways induced by Rho-family GTPases. Therefore, we tested the effects of expression of a kinase-inactive form of LIMK-1 on RhoV14-, RacV12- and Cdc42V12-induced actin reorganization. Expression of LIMK-1(D460A) significantly repressed RacV12-induced lamellipodium formation (Fig. 3a); lamellipodium formation occurred in 90% of RacV12-injected cells (62 out of 69 RacV12-positive cells), but only in 55% of cells injected with RacV12 and LIMK-1(D460A) (32 out of 58 RacV12-positive cells). In contrast, expression of LIMK-1(D460A) had no apparent effect on RhoV14-induced stress fibre formation or on Cdc42V12-induced actin reorganization (Fig. 3a). Thus, LIMK-1 is implicated in Rac-induced lamellipodium formation. Some increase in the number of stress fibres was consistently observed in cells injected with RacV12 and LIMK-1(D460A) (Fig. 3a), indicating that RacV12 may stimulate Rho in these cells, as seen in fibroblasts³.

Next we determined whether an active form of Rac could stimulate the kinase activity of LIMK-1. We transfected LIMK-1 with RacV12; the kinase activity of LIMK-1 then increased about twofold, as measured by an *in vitro* kinase reaction using (His)₆-cofilin as a substrate (Fig. 3b). In contrast, RhoV14 had no effect and

Cdc42V12 slightly enhanced LIMK-1 activity. As Cdc42 is known to stimulate Rac activity⁴, Cdc42V12-induced activation of LIMK-1 may occur through the activation of Rac. These results indicate that LIMK-1 is a putative downstream effector of Rac and may be involved in Rac-induced lamellipodium formation. However, we saw neither a direct interaction between LIMK-1 and Rac (either GTP- or GDP-bound) nor *in vitro* activation of LIMK-1 by GTP-bound Rac; therefore, the observed Rac-induced activation of LIMK-1 *in vivo* is indirect.

Insulin induces the formation of lamellipodia and membrane ruffings in several types of cell by a process that depends on Rac^{3,5}. Consistent with earlier observations⁵, insulin-induced membrane ruffling in HeLa cells (data not shown) and KB cells (Fig. 4b) was detectable within 2 min, reached a maximum at 5–10 min, and decreased within 30 min of exposure to insulin. We examined changes in the kinase activity of endogenous LIMK-1 in HeLa cells after insulin treatment. Kinase activity of LIMK-1, as measured by the *in vitro* kinase assay, increased 1.3- to 1.4-fold at 5–10 min after the addition of insulin and then decreased gradually to the original level (Fig. 4a). The time course of LIMK-1 activation almost paralleled that of insulin-stimulated formation of membrane ruf-

fls, indicating a role for LIMK-1 activation in ruffle formation. When we examined the level of phosphorylation of endogenous cofilin in HeLa cells before and after treatment with insulin, we saw no detectable change. It could be that, under conditions in which the cell is ruffling, while cofilin is phosphorylated by activated LIMK-1 at specific sites, it may also be dephosphorylated at other sites.

To determine whether endogenous LIMK-1 is involved in the process of insulin-induced formation of membrane ruffles, we transfected KB cells with GFP–LIMK1(D460A) (a kinase inactive LIMK-1 fused with green fluorescence protein (GFP)) or with control GFP, and then treated these cells with insulin. The insulin-induced membrane ruffling was significantly suppressed in cells expressing GFP–LIMK1(D460A), but not in control, GFP-expressing cells (Fig. 4b); insulin treatment induced ruffling in 76% (71 of 93) of GFP-expressing cells, but only in 42% (36 of 85) of cells expressing GFP–LIMK1(D460A). These findings indicate that endogenous LIMK-1 participates in insulin-induced membrane ruffle formation.

We propose that LIMK-1 is involved in signalling pathways of stimulus-induced actin reorganization. Our results indicate a

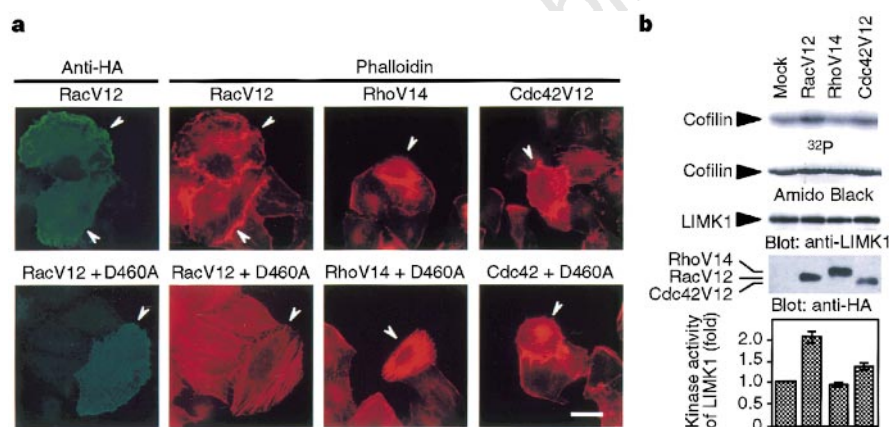


Figure 3 Involvement of LIMK-1 in Rac signalling. **a**, LIMK-1(D460A) inhibits Rac-induced lamellipodium formation. HeLa cells injected with plasmids encoding haemagglutinin (HA)-tagged RacV12, RhoV14 or Cdc42V12, with or without plasmid for Flag-tagged LIMK-1(D460A), were stained with anti-HA antibody and with rhodamine-phalloidin. Arrowheads indicate the injected cells, as assigned by anti-HA immunostaining. Scale bar, 20 μm. **b**, RacV12 increases LIMK-1

activity. COS cells were transfected with only plasmid for LIMK-1 (mock) or with this plasmid and plasmids encoding HA-tagged RacV12, RhoV14 or Cdc42V12 (as indicated). LIMK-1 was immunoprecipitated and subjected to *in vitro* kinase assay using (His)₆-cofilin as a substrate. The specific activity of LIMK-1 was calculated (see Methods). Results are shown as means ± s.e.m. of triplicate experiments, taking the value for mock cells as 1.0.

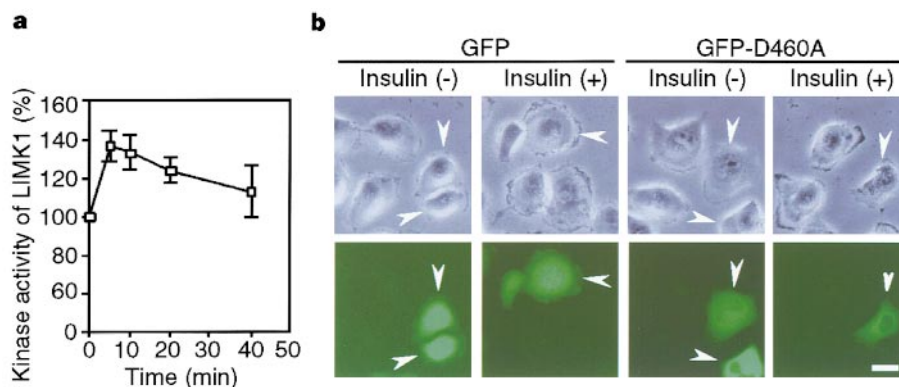


Figure 4 Involvement of LIMK-1 in insulin-induced formation of membrane ruffling. **a**, Insulin-induced activation of LIMK-1. Endogenous LIMK-1 in HeLa cells was immunoprecipitated at the indicated times after treatment with insulin (10 μg ml⁻¹), and was subjected to the *in vitro* kinase assay, using (His)₆-cofilin as a substrate. Specific activity was calculated (see Methods). Results are shown as means ± s.e.m. from triplicate experiments, taking the value at time 0 as 100%. **b**, Inhibition of insulin-induced membrane ruffings by expression of

kinase-inactive LIMK-1. KB cells transfected with plasmid encoding GFP or GFP–LIMK1(D460A) were stimulated with insulin (10 μg ml⁻¹) for 10 min. Membrane ruffings were visualized as phase-dark regions at cell margins by phase-contrast microscopy (upper panels). GFP expression was visualized by fluorescence microscopy (lower panels). Arrowheads indicate GFP-positive cells. Scale bar, 20 μm.

potential link between Rac, LIMK-1, cofilin and actin-filament dynamics; Rac stimulates the kinase activity of LIMK-1, which induces phosphorylation and inactivation of cofilin and probably leads to a decrease in the rate of actin depolymerization. The binding of LIMK-1 to F-actin may localize LIMK-1 activity to confined regions within the cell. The LIM and PDZ domains of LIMK-1 may also help to define its localization. As cofilin is a potent and crucial regulator of actin-filament dynamics^{6–10}, our finding that LIMK-1 is a cofilin-phosphorylating enzyme will be important in understanding the signalling mechanisms that regulate actin filament dynamics. □

Methods

Plasmids, antibody and GST fusion proteins. Expression plasmids encoding LIMK-1, cofilin and their mutants were constructed in pUC-SRα, pGEX-2T and pQE60 vectors, as described^{12,15,18}. Plasmids for Rho family GTPases in pEF-Bos vector¹⁹ were provided by S. Kuroda, K. Kaibuchi and K. Ito. The plasmid for GFP-LIMK1(D460A) was constructed by inserting the cDNA of LIMK1(D460A)(215–647) into pEGFP-C1 (Clontech). Anti-LIMK1 antibody (LK-C10) was prepared as described¹⁵. GST fusion proteins were prepared as described^{12,18}.

Purification of LIMK1-binding proteins. We homogenized rat brains (5.4 g) in 22 ml buffer A (50 mM Tris-HCl buffer, pH 8.0, 100 mM NaCl, 1 mM phenylmethylsulphonyl fluoride). After addition of NP-40 to 0.1%, homogenates were incubated for 30 min on ice and centrifuged at 36,000g for 40 min. The supernatant was incubated with GST-coupled glutathione-Sepharose (0.5 ml), and the non-adsorbed fraction was loaded on a GST-LIMK1- or GST-coupled glutathione-Sepharose column (0.5 ml). The column was washed with buffer A containing 0.1% NP-40, and the bound proteins were eluted with 50 mM Tris-HCl buffer (pH 8.0) containing 1 M NaCl and 0.1% NP-40, run on SDS-PAGE, and transferred onto a polyvinylidene difluoride membrane. The membrane containing p42 was excised and digested with lysyl endopeptidase. The peptides were separated by HPLC (μBondasphere C-18, Waters). The sequence of each peptide was analysed using a protein sequencer 494 (Applied Biosystems). The membrane was also analyzed by immunoblots, using an anti-β-actin antibody (Biomedical Technology).

F-actin-binding assay. The F-actin-binding assay was performed as described²⁰. GST fusion proteins expressed in *Escherichia coli* were incubated with F-actin for 2 h at 4 °C. The reaction mixture was centrifuged at 56,000g for 20 min, and equal portions of the supernatant and pellets were subjected to SDS-PAGE and were then analysed by immunoblotting, using an anti-GST or anti-LIMK1 antibody.

Transfection, microinjection and staining. Cells were plated on glass coverslips, and plasmids were transfected by lipofectamine (Gibco) or calcium phosphate precipitation methods or were microinjected into the nucleus, as described^{15,19}. After 16–48 h, cells were fixed and immunostained using M2 anti-Flag (Kodak) or 12CA5 anti-haemagglutinin monoclonal antibody (Babco), followed by fluorescein-isothiocyanate-labelled secondary antibody (Amersham), as described¹⁵. Cells were also stained for F-actin with rhodamine-phalloidin.

In vitro kinase assay. Lysates prepared from cultured cells were immunoprecipitated with anti-LIMK1 antibody¹⁵, and included in an *in vitro* kinase reaction with 50 μM ATP, 5 μCi of [γ -³²P]ATP (3,000 Ci mmol⁻¹) and (His)₆-cofilin (0.25 mg ml⁻¹), as described¹⁵. Specific activity of LIMK-1 was calculated by dividing ³²P-radioactivity incorporated into (His)₆-cofilin by the density of anti-LIMK1 immunoreactive band.

In vivo kinase assay. COS-7 cells were transfected with plasmids encoding C-terminally Sky-peptide-tagged²¹ cofilin and LIMK-1. Cells were labelled with [³²P]phosphoric acid, as described¹². Lysates were immunoprecipitated with anti-Sky antibody²¹, and immunoprecipitates were run on SDS-PAGE and analysed by autoradiography and immunoblotting with anti-Sky antibody.

Received 10 March; accepted 15 April 1998.

- Hall, A. Rho GTPases and the actin cytoskeleton. *Science* **279**, 509–514 (1998).
- Tarpon, N. & Hall, A. Rho, Rac and Cdc42 GTPases regulate the organization of the actin cytoskeleton. *Curr. Opin. Cell Biol.* **9**, 86–92 (1997).
- Ridley, A. J., Paterson, H. F., Johnston, C. L., Diekmann, D. & Hall, A. The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell* **70**, 401–410 (1992).

- Nobes, C. D. & Hall, A. Rho, Rac, and Cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* **81**, 53–62 (1995).
- Nishiyama, T. *et al.* rac p21 is involved in insulin-induced membrane ruffling and rho p21 is involved in hepatocyte growth factor- and 12-O-tetradecanoylphorbol-13-acetate (TPA)-induced membrane ruffling in KB cells. *Mol. Cell. Biol.* **14**, 2447–2456 (1994).
- Nishida, E., Maekawa, S. & Sakai, H. Cofilin, a protein in porcine brain that binds to actin filaments and inhibits their interactions with myosin and tropomyosin. *Biochemistry* **23**, 5307–5313 (1984).
- Moon, A. & Drubin, D. G. The ADF/cofilin proteins: stimulus-responsive modulators of actin dynamics. *Mol. Biol. Cell* **6**, 1423–1431 (1995).
- Therriot, J. A. Accelerating on a treadmill: ADF/cofilin promotes rapid actin filament turnover in the dynamic cytoskeleton. *J. Cell Biol.* **136**, 1165–1168 (1997).
- Welch, M. D., Mallavarapu, A., Rosenblatt, J. & Mitchison, T. J. Actin dynamics *in vivo*. *Curr. Opin. Cell Biol.* **9**, 54–61 (1997).
- Carlier, M.-F. & Pantaloni, D. Control of actin dynamics in cell motility. *J. Mol. Biol.* **269**, 459–467 (1997).
- Agnew, B. J., Minamide, L. S. & Bamberg, J. R. Reactivation of phosphorylated actin depolymerizing factor and identification of the regulatory site. *J. Biol. Chem.* **270**, 17582–17587 (1995).
- Moriyama, K., Iida, K. & Yahara, I. Phosphorylation of Ser-3 of cofilin regulates its essential function on actin. *Genes Cells* **1**, 73–86 (1996).
- Mizuno, K. *et al.* Identification of a human cDNA encoding a novel protein kinase with two repeats of the LIM/double zinc finger motif. *Oncogene* **9**, 1605–1612 (1994).
- Bernard, O., Ganiatsas, S., Kannourakis, G. & Dringen, R. Kiz-1, a protein with LIM zinc finger and kinase domains, is expressed mainly in neurons. *Cell Growth Differ.* **5**, 1159–1171 (1994).
- Okano, I. *et al.* Identification and characterization of a novel family of serine/threonine kinases containing N-terminal two LIM motifs. *J. Biol. Chem.* **270**, 31321–31330 (1995).
- Nunoue, K., Ohashi, K., Okano, I. & Mizuno, K. LIMK-1 and LIMK-2, two members of a LIM motif-containing protein kinase family. *Oncogene* **11**, 701–710 (1995).
- Ridley, A. J. & Hall, A. The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* **70**, 389–399 (1992).
- Hiraoka, J., Okano, I., Higuchi, O., Yang, N. & Mizuno, K. Self-association of LIM-kinase 1 mediated by the interaction between an N-terminal LIM domain and a C-terminal kinase domain. *FEBS Lett.* **399**, 117–121 (1996).
- Kuroda, S. *et al.* Identification of IQGAP as a putative target for the small GTPases, CDC42 and Rac1. *J. Biol. Chem.* **271**, 23363–23367 (1996).
- Roof, D. J., Hayes, A., Adamian, M., Chishti, A. H. & Li, T. Molecular characterization of abLIM, a novel actin-binding and double zinc finger protein. *J. Cell Biol.* **138**, 575–588 (1997).
- Ohashi, K. *et al.* Stimulation of Sky receptor tyrosine kinase by the product of growth arrest-specific gene 6. *J. Biol. Chem.* **270**, 22681–22684 (1995).

Acknowledgements. We thank S. Kuroda, K. Kaibuchi and K. Ito for plasmids of Rho, Rac and Cdc42; T. Amano and A. Miyata for technical assistance; and Y. Fujiki, H. Matsuo, C. Ikebe, M. Ohara and K. Toyoshima for comments. This work was supported by grants from the Ministry of Education, Science, Sports and Culture of Japan and the Japan Science and Technology Corporation (to K.M.).

Correspondence and requests for materials should be addressed to K.M. (e-mail: kmizusc@mbn.nc.kyushu-u.ac.jp).

Structure of a heparin-linked biologically active dimer of fibroblast growth factor

Anna D. DiGabriele*, Irit Lax†, Denise I. Chen*, Carl M. Svahn‡, Michael Jaye§, Joseph Schlessinger† & Wayne A. Hendrickson*||

* Department of Biochemistry and Molecular Biophysics, and || Howard Hughes Medical Institute, Columbia University, New York, New York 10032, USA

† Department of Pharmacology, New York University Medical Center, New York, New York 10016, USA

‡ Pharmacia & Upjohn, Lindhagensgatan 133, S-11287 Stockholm, Sweden

§ Rhône-Poulenc-Rorer, 500 Arcola Road, Collegeville, Pennsylvania 19426, USA

The fibroblast growth factors (FGFs) form a large family of structurally related, multifunctional proteins that regulate various biological responses¹. They mediate cellular functions by binding to transmembrane FGF receptors², which are protein tyrosine kinases. FGF receptors are activated by oligomerization³, and both this activation and FGF-stimulated biological responses require heparin-like molecules as well as FGF⁴. Heparins are linear anionic polysaccharide chains; they are typically heterogeneously sulphated on alternating L-iduronic and D-glucosamine sugars, and are nearly ubiquitous in animal tissues as heparan sulphate proteoglycans on cell surfaces and in the extracellular matrix. Although several crystal structures have been described for FGF molecules in complexes with heparin-like sugars^{5–7}, the nature of a biologically active complex has been unknown until now. Here we describe the X-ray crystal structure, at 2.9 Å

resolution, of a biologically active dimer of human acidic FGF in a complex with a fully sulphated, homogeneous heparin decasaccharide. The dimerization of heparin-linked acidic FGF observed here is an elegant mechanism for the modulation of signalling through combinatorial homodimerization and heterodimerization of the 12 known members of the FGF family.

We examined the decasaccharide used in the crystallization studies for its ability to oligomerize acidic FGF (aFGF), and tested the complex for biological activity in living cells. The sugar dimerizes aFGF in solution, as shown by all size-determination techniques used here, including gel filtration, dynamic light scattering, and sedimentation velocity ultracentrifugation (Table 1). Furthermore, the aFGF-heparin complex promotes tyrosine autophosphorylation and activation of the FGF receptor (FGFR) and induces neurite outgrowth in living cells (Fig. 1).

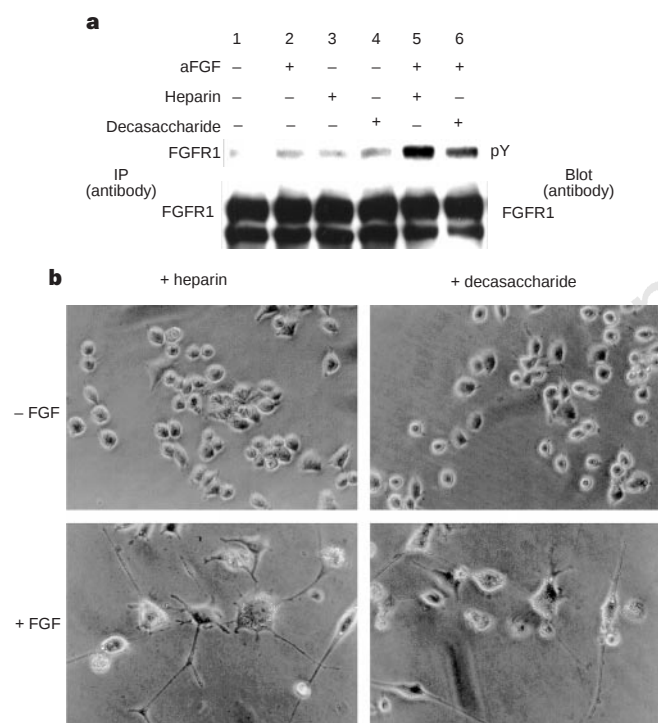


Figure 1 Biological activity of aFGF-heparin and aFGF-decasaccharide complexes. **a**, Activation of FGFR1 by the aFGF-decasaccharide complex. BaF3 cells expressing FGFR1 and the indicated combinations of aFGF, heparin and decasaccharide were lysed, and FGFR1 was immunoprecipitated (IP) with anti-FGFR1 antibodies. Samples were analysed by SDS-PAGE, transferred to a nitrocellulose membrane and subjected to immunoblotting with either anti-phosphotyrosine (pY) or anti-FGFR antibodies. **b**, The aFGF-decasaccharide complex promotes neurite outgrowth in PC12 cells. PC12 cells, stably transfected with FGFR-1, were left untreated or were stimulated with the indicated combinations of aFGF, decasaccharide and heparin. Neurite outgrowth was observed 24 h after ligand treatment. Cells in the absence of both aFGF and sugar resemble untreated cells (data not shown).

Complexes of the decasaccharide with both native and selenomethionyl aFGF were purified and crystallized. Molecular replacement gave an initial structure for a hexagonal crystal form; multiwavelength anomalous diffraction (MAD) analysis of a related orthorhombic crystal form was also used because electron density for the carbohydrate was difficult to interpret (Table 2). Hexose units of each heparin chain could be positioned into the resulting electron-density map (Fig. 2a). The native aFGF/heparin and selenomethionyl aFGF/heparin structures were refined at 2.9 Å and 3.0 Å, respectively (Table 2). The two crystal forms each contain a total of four independent heparin-linked aFGF dimers. Heparin adopts a 1C_4 helical conformation and is bound to aFGF-dimer mates in two distinct orientations. In the structures of all four dimers, five (orthorhombic) or six (hexagonal) of the ten hexose units of the decasaccharide chains are ordered and bind through sulphate groups to residues in the heparin-binding loop of aFGF and heparin (residues 112–128).

The crystal structure of the biologically active, heparin-decasaccharide-linked dimer of aFGF exhibits a unique mode of ligand-induced protein dimerization. This aFGF dimer is bridged by heparin and does not have a protein-protein interface (Fig. 2b, c). Instead, the protomers are positioned to bind sulphate groups of 5–6 monosaccharide units on opposite sides of the heparin helix axis. A quasi-dyad axis that relates the two aFGF protomers passes perpendicularly through the heparin helix axis (Fig. 2b). This axis of symmetry is not a true two-fold axis because of the sugar polarity. Heparin binding does not alter aFGF conformation significantly (root-mean-square (r.m.s.) deviations in $C\alpha$ between the eight protomers in the crystal forms described here and unliganded aFGF (2AFG) are 0.40–0.55 Å), in agreement with earlier descriptions of crystal structures^{5–7}. Similarly, the decasaccharide structures observed here, the solution conformation of a heparin dodecasaccharide⁸, and the structures of the tetrasaccharide and hexasaccharides as bound in basic FGF (bFGF) crystals⁷, are nearly identical.

Comparison of the four independent heparin-linked dimers in the two crystal forms shows that there is appreciable flexibility in the heparin-mediated union (Fig. 2c). The aFGF protomers in dimers A, B and H bind to structurally equivalent parts of decasaccharide chains A_{sugar} , B_{sugar} and H_{sugar} respectively, and are therefore in similar orientations with respect to each other. The aFGF protomers in dimer C have C_{sugar} bound with the chain displaced by one monosaccharide unit. As the structural repeat of heparin is four hexose units, the sulphate groups along the heparin chain are in slightly different positions when the binding site is shifted by one monosaccharide unit, and therefore the C5 and C6 protomers are in different positions with respect to the A1 and A2 protomers.

The heparin polysaccharide is polar, with one end (O1) reducing and the other end (O4) non-reducing. Despite this polarity, the disposition of heparin sulphate groups on either side of the helix axis is such that the quasi-dyad axis that relates aFGF protomers of one heparin-linked dimer is accommodated by the binding of decasaccharide oppositely to the aFGF-dimer mates (Fig. 3a–d). Although the decasaccharide chain is positioned, with respect to the sugar-binding loop, very similarly to the oligosaccharide chains in

Table 1 Determination of relative molecular masses of aFGF and aFGF-oligosaccharide complexes

	Superose-12 gel filtration		Dynamic light scattering		Ultracentrifugation	
	Vol (ml)	M_r (K)	D ($\times 10^{-7}$ cm ² s ⁻¹)	M_r (K)	S ($\times 10^{-13}$ s)	M_r (K)
aFGF	14.7	16.8	10.9	17.8	3.0	23.4
aFGF-SOS (1.0K)	14.2	—	10.5	19.9	—	—
aFGF-hexasaccharide (1.8K)	13.8	26.9	10.9	18.5	—	—
aFGF-octasaccharide (2.5K)	13.1	35.9	9.1	29.1	—	—
aFGF-decasaccharide (3.0K)	12.9	40.5	8.4	33.0	3.4	36.0

Approximate relative molecular masses of oligosaccharides are in parentheses in the first column. Data are averages from multiple experiments. The results are compatible with dimer formation for octa- and decasaccharide complexes. D , translational diffusion coefficient; S , sedimentation coefficient.

the bFGF–oligosaccharide structures⁷, the shorter sugars are bound to bFGF in only one orientation (O1 to O4) (Fig. 3a, c).

The sugar-binding site on aFGF is formed by Asn 18 and by side chains of the surface loop that consists of residues 112–128. This is the same or analogous region that binds sucrose octasulphate (SOS)⁵ and sulphated and unsulphated oligosaccharides^{6,7} in other crystal structures. Most of the interactions between aFGF and the decasaccharide are ionic contacts between basic residues and the sulphate and carboxylate groups of heparin. Details of decasaccharide–aFGF interactions for protomers in dimers A and C (Fig. 3a–d) indicate that no two interactions are exactly alike. Thus, Lys 118 and, differently in the eight interfaces, one or two amino acids from among Lys 112, Lys 113, Arg 122 and Lys 128 form hydrogen bonds with sugar sulphate or carboxylate groups.

The variations in heparin-linked aFGF–dimer structures and in aFGF protomer/sugar interactions seen here (Figs 2c and 3a–d) may reflect the biological diversity of interactions. As the heparin sulphates associated with target cells are heterogeneously sulphated, FGF molecules must bind different heparin-like sugars to produce biological functions. The long side chains of basic residues in the binding site provide conformational flexibility to accommodate differences in sugar sulphation and alternative sugar directions.

Despite variability in details, the FGF–heparin interactions are remarkably similar, given that the paths of the heparin helices are distinct when bound in opposite orientations and that the heparin chain is offset in dimer C. A least-squares superposition of all eight aFGF protomers in our two lattices, together with their carbohydrate ligands, shows the underlying similarity in sugar binding (Fig. 3e). There are three principal sulphate-group-binding sites within the FGF sugar-binding loop and, despite the differences in heparin helix path and orientation, either two or three sulphate groups from each chain bind at these sites. Furthermore, the side chains of aFGF amino-acid residues forming the three sites are in similar conformations in all of the FGF protomers. In all eight aFGF protomers, sulphate groups occupy sites 1 and 2, which are formed

from the side chains of residues Asn 18, Lys 118 and Gln 127 (Fig. 3e). In addition, binding site 3, which is formed from the side chain of Asn 114 and the peptide backbone of residues 113 and 114, is occupied in each of the four protomers (A2, B4, C6 and H8) that bind the decasaccharide in the O4 to O1 orientation (Fig. 3b, d, e). The sugar-binding role of these residues is corroborated by structural^{5–7} and mutagenesis^{9,10} studies.

Although it has been established that one aFGF contacts 4–5 monosaccharide units⁴ and we find only 5–6 monosaccharide units ordered in our crystals, a minimum of eight monosaccharide units is required to form a biologically active FGF dimer¹¹. One plausible explanation for such a length requirement is provided by study of electrostatics. The sugar-binding loop of aFGF is a region of large, positive electrostatic potential (Fig. 3f), and the dimerized sugar-binding sites (12 positive charges per aFGF dimer) are poised to repel (Figs 2 and 3f). Only four to five of these charges are directly neutralized by salt bridges in the heparin complex; thus, even a fully sulphated hexasaccharide (12 negative charges) may not be enough to counteract the repulsion as the more remotely placed charges contribute little to the electrostatic free energy. Alternative explanations for the apparent discrepancy may be the existence of additional low-affinity sugar–FGF contacts that are insufficiently occupied to be observed, that are bound but conformationally heterogeneous, or that are displaced by the forces of crystal packing.

There is no evidence here of aFGF dimerization by SOS or by a heparin tetrasaccharide or hexasaccharide (Table 1)⁴. In contrast, results of studies of basic FGF and heparin oligosaccharides by either dynamic light scattering and NMR¹² or by analytical ultracentrifugation¹³ have been interpreted to indicate that a sugar as small as SOS or a tetrasaccharide can induce FGF dimerization. Moreover, it has been proposed that, as an octasaccharide is the smallest biologically active heparin unit, any FGF dimer induced to form by a shorter oligosaccharide must be of a so-called ‘sandwich’ form¹³, not unlike the structures described here. Instead, it was suggested that the active dimer consists of two bFGF molecules

Table 2 Crystallographic analysis

Diffraction data						MAD structure factor ratio†‡				
Data set (space group)	Wavelength (Å)	Resolution (Å)	Reflections*	Completeness (%)	$R_{\text{sym}}^{\dagger}$ (%)	30.0 Å < d < 5.0 Å				
Native (P6 ₂ 2 ₂)	0.98008	20–2.8	12,31 (132,305)	99.6	21.0	Wavelength (Å)	λ_1	λ_2	λ_3	λ_4
Se-met MAD (C222 ₁)	0.98789	30–2.8	112,646 (123,799)	85.1	9.6	λ_1 (0.98789)	0.055 (0.048)	0.054	0.049	0.045
	0.97930	30–2.8	101,323 (120,478)	83.4	9.6	λ_2 (0.97930)	–	0.062 (0.051)	0.047	0.056
	0.97907	30–2.8	106,382 (115,940)	83.0	9.6	λ_3 (0.97907)	–	–	0.073 (0.054)	0.051
	0.96859	30–2.8	108,962 (119,439)	83.6	9.3	λ_4 (0.96859)	Number of sites = 6 ($\langle I^2/F_1 \rangle$) = 0.057 $R(\langle I^2/F_2 \rangle)$ = 0.517 (FOM)§ = 0.69			
Refinement space group	Resolution (Å)	R-value¶ (Test R-value)	r.m.s. bond distance (Å)	r.m.s. angles (°)	B-values (Å ²)	Model number of non-H atoms (number of waters)	Number of reflections (number in test set)	Ramachandran plot: residues in generously allowed regions (%)		
P6 ₂ 2 ₂	13–2.9	21.8 (30.6)	0.013	1.73	28.4	2,159 (13)	8,818 (1,007)	0.9		
C222 ₁	13–3.0	21.7 (30.4)	0.014	1.63	37.6	6,399 (71)	21,196 (2,344)	2.8		

* For native data set, unique reflections counted in point group 622. For MAD data set, unique acentric reflections were counted in point group 222 and unique centric reflections were counted in point group mmm. Values in parentheses are total number of measurements.

† $R_{\text{sym}} = 100 \times \sum_i \sum_j |I_i - \langle I \rangle| / \sum_i \sum_j I_i$, where I_i is the i th measurement, $\langle I \rangle$ is the weighted mean of all measurements of I , and h, k, l are the reflection indices.

‡ Values for structure factor ratios are r.m.s. ($\Delta|F|$)/r.m.s. ($|F|$) where $\Delta|F|$ is the Bijvoet difference at a single wavelength (diagonal) or dispersive differences between two wavelengths (off-diagonal). Values in parentheses are ratios for centric Bijvoet reflections and would be zero for perfect data and are an estimate of noise by which to gauge anomalous signals.

§ (FOM) is the mean figure of merit including reflections with FOM = 0.

|| Statistics for data with $F > 2\sigma$.

¶ R-value = $100 \times \sum_i ||F_o| - |F_c|| / \sum_i |F_o|$, where F_o are the observed structure factors and F_c are the calculated structure factors.

bound to the oligosaccharide on the same face of a single sugar, and sharing a protein–protein interface¹³. There is, however, no tetrasaccharide or hexasaccharide-induced ‘sandwich’ dimer of bFGF in reported complex crystal structures⁷, where concentrations of protein and sugar are greater than here and an examination of various oligosaccharide-bound and apo-FGF structures shows that no protein–protein contact is conserved in the different crystal lattices.

By combining crystallographic and mutagenesis data, we can propose a model for the dimerization of FGFRs that may explain how the FGF ligands transmit different types of signal in many different cells expressing the various isoforms of FGFRs. The FGF residues that were suggested by mutagenesis experiments^{14,15} to be important for FGFR binding were mapped onto the molecular surface of the heparin-linked aFGF dimer (Fig. 3g). This map shows that these FGF residues cluster along opposite faces of each protomer. The distance between these FGFR-binding sites is ~ 40 Å,

contrasting with the distance of ~ 40 – 60 Å that is found between Flt-1-binding sites on the vascular endothelial growth factor (VEGF) dimer in the Flt-1/VEGF complex¹⁶. The model for dimerization of protein tyrosine kinases derived from our data (Fig. 3h) is distinguished from earlier models^{4,13,15} by a lack of both FGF–FGF and heparin–receptor interfaces. The existence of many FGF subtypes and multiple FGFR isoforms indicates that heterodimerization of FGFs and of FGFRs is partially responsible for the variety of FGF activities. The plasticity that we observe in heparin–FGF interfaces would facilitate such FGF heterodimerization. Validation of the model (Fig. 3h) awaits structural analysis of an appropriate FGF–heparin–FGFR complex. □

Methods

Solution and biological studies. Human aFGF formed complexes with heparin oligosaccharides in a 2:1 molar ratio. Hexasaccharides and octasaccharides

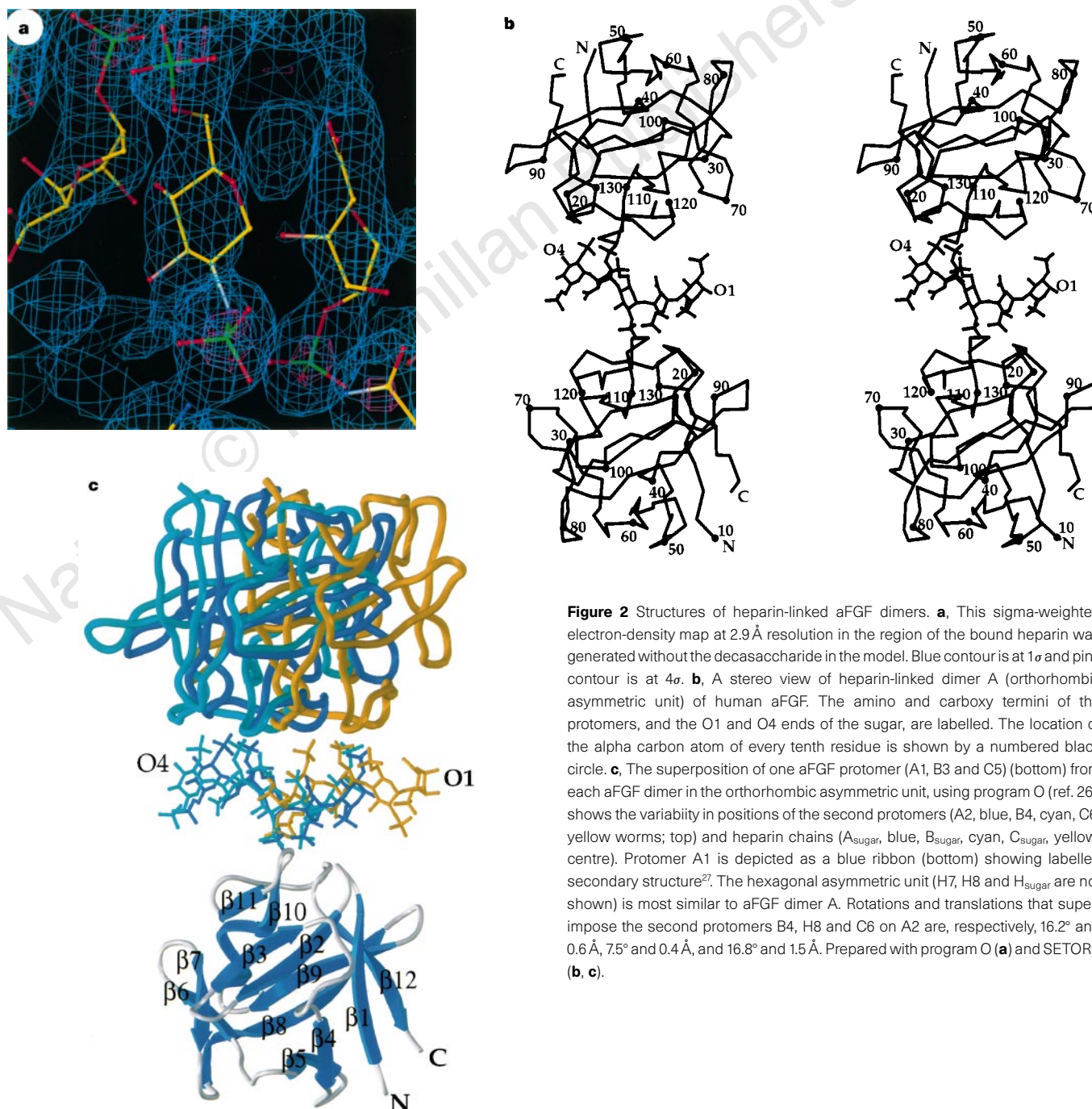


Figure 2 Structures of heparin-linked aFGF dimers. **a**, This sigma-weighted electron-density map at 2.9 Å resolution in the region of the bound heparin was generated without the deca-saccharide in the model. Blue contour is at 1σ and pink contour is at 4σ . **b**, A stereo view of heparin-linked dimer A (orthorhombic asymmetric unit) of human aFGF. The amino and carboxy termini of the protomers, and the O1 and O4 ends of the sugar, are labelled. The location of the alpha carbon atom of every tenth residue is shown by a numbered black circle. **c**, The superposition of one aFGF protomer (A1, B3 and C5) (bottom) from each aFGF dimer in the orthorhombic asymmetric unit, using program O (ref. 26), shows the variability in positions of the second protomers (A2, blue, B4, cyan, C6, yellow worms; top) and heparin chains (A_{sugar}, blue, B_{sugar}, cyan, C_{sugar}, yellow; centre). Protomer A1 is depicted as a blue ribbon (bottom) showing labelled secondary structure²⁷. The hexagonal asymmetric unit (H7, H8 and H_{sugar} are not shown) is most similar to aFGF dimer A. Rotations and translations that superimpose the second protomers B4, H8 and C6 on A2 are, respectively, 16.2° and 0.6 Å, 7.5° and 0.4 Å, and 16.8° and 1.5 Å. Prepared with program O (**a**) and SETOR²⁸ (**b**, **c**).

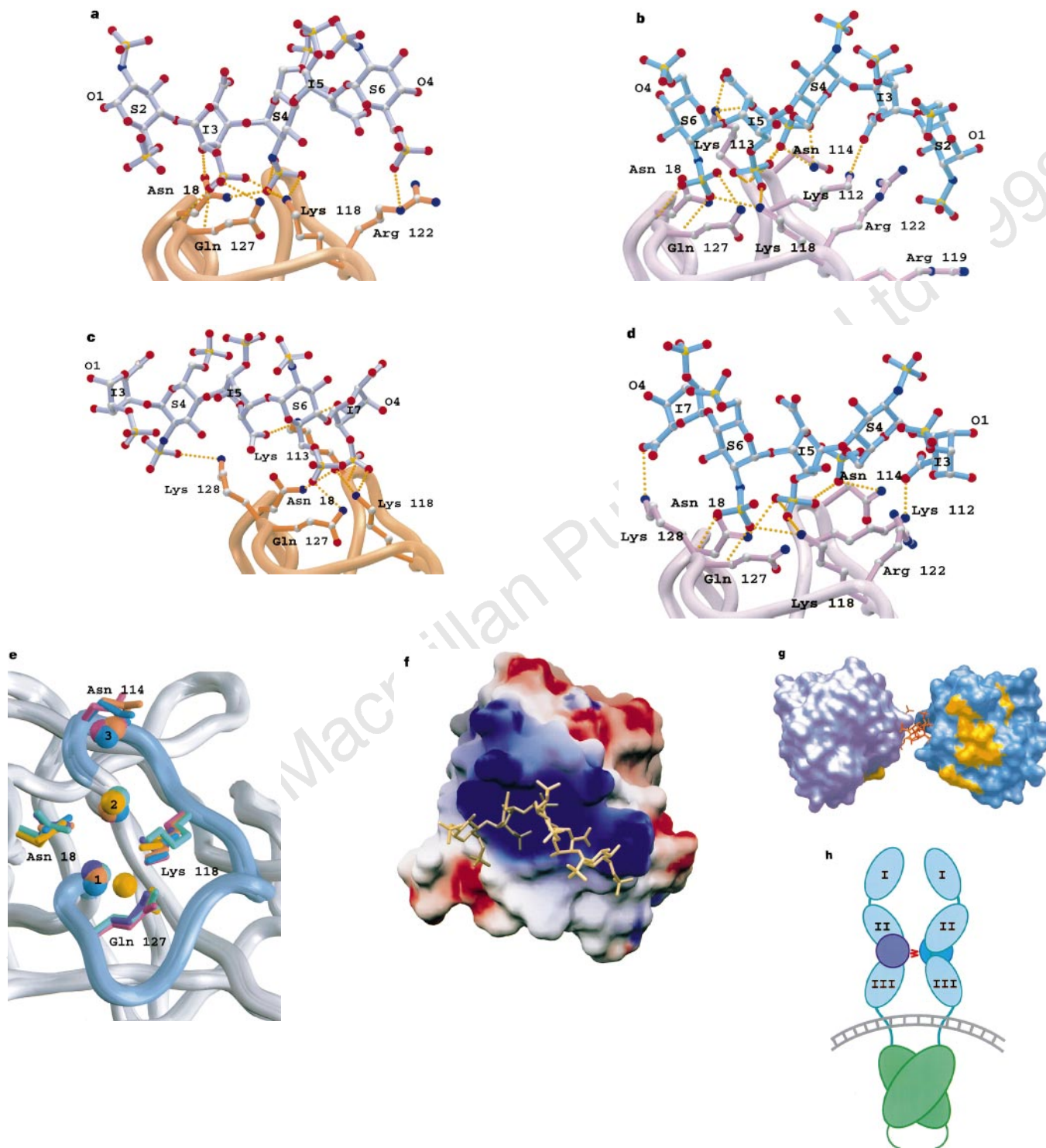


Figure 3 a-d, Details of hydrogen bonding between aFGF protomers (ribbons) and heparin (ball-and-stick diagrams) in aFGF dimers A and C. After C α superposition, protomers of one dimer and bound sugar were placed side by side for comparison. A1 (**a**) and C5 (**c**) bind A_{sugar} and C_{sugar}, respectively, with 'O1 to O4' polarity. A2 (**b**) and C6 (**d**) bind the corresponding sugars with 'O4 to O1' polarity. All side chains within 4 Å of the sugar are shown. Atoms are colour-coded by element (C, white; O, red; S, yellow; N, blue) and interactions are shown as dashed yellow lines. Monosaccharide units are labelled as S for N-acetyl glucosamine and I for iduronic acid. **e**, Superposition of the six aFGF protomers and bound heparin chains in the orthorhombic asymmetric unit looking down on the sugar-binding loop (blue worm segment). Protein side chains and corresponding sulphate groups (spheres) are a different colour for each protomer.

Three major sulphate-group binding sites and the amino-acid side chains that form these sites are shown. **f**, Electrostatic potential mapped onto the molecular surface of an aFGF protomer orientated as in **e**. A large patch of positive electrostatic potential (blue represents +10 e per Å) distinguishes the heparin-binding site (heparin is shown in yellow). **g**, Surface representation of aFGF dimer A (Fig. 2b) rotated by 90° about two perpendicular axes. Yellow surfaces represent FGFR-binding sites. A two-fold axis runs vertically in the plane of the page between the protomers, such that the yellow surface at the front of the blue protomer is at the back of the purple protomer. **h**, A model for FGFR dimerization by binding of the extracellular (numbered) FGFR domains (cyan) to the heparin (orange zigzag)-linked aFGF dimer (blue and purple circles), orientated as in **g**. Prepared with SETOR (**a-d**, **e**, **g**)²⁸ and GRASP (**f**)²⁹.

were obtained from Novo Nordisk, and a fully sulphated decasaccharide (COA30723) was fractionated at Pharmacia from natural heparin. Gel filtration was done using a calibrated Superose-12 column on an FPLC system (Pharmacia). Dynamic light-scattering data were collected for multiple scans and translational diffusion coefficients were calculated (DynaPro-801 Molecular Size Detector, ProteinSolutions). Sedimentation velocity experiments (UTHSCSA) were done at room temperature and 42,000 r.p.m. using a Beckman XL-A analytical ultracentrifuge equipped with scanning absorption optics. We used the vanHolde-Weischet method of analysis¹⁷ to determine sedimentation coefficients.

BaF3 and PC12 cells overexpressing FGFR-1 have been described^{4,18}. We generated anti-FGFR1 antibodies against a glutathione *S*-transferase (GST) fusion protein composed of the carboxy-terminal region of FGFR-1. Anti-phosphotyrosine antibodies were generated as described⁴. Interleukin-3 (IL-3)-dependent BaF3 cells were starved for 6–8 h in medium containing 10% FBS to remove IL-3. The cells were incubated for 10 min at 37 °C with aFGF (0.6 nM) and either heparin (0.6 µM) or decasaccharide (0.6 µM). The cells were then lysed as described⁴, immunoprecipitated with anti-FGFR1 antibodies, and analysed by immunoblotting with either anti-phosphotyrosine antibodies or anti-FGFR1 antibodies. PC12 cells were seeded in 6-cm tissue-culture plates at a density of 2×10^4 cells per plate. The cells were incubated with complete growth medium to which aFGF (4 nM) and either heparin (0.6 µM) or decasaccharide (0.6 µM) were added separately or together.

Crystallization. Native and selenomethionyl-aFGF were produced and purified from *Escherichia coli*^{19,20}. Each protein was complexed with heparin decasaccharide (COA30723, Pharmacia) in a molar ratio of 2:1. Complexes were crystallized at 20 °C by vapour diffusion in hanging drops containing equal volumes of 20 mg ml⁻¹ protein-sugar solution and the reservoir solution (containing 20% PEG 8000, 200 mM MgSO₄ and 100 mM HEPES, pH 7.0). Semet-aFGF/sugar complex crystallizations were done with the addition of 20 mM dithiothreitol and drops were equilibrated in an oxygen-free atmosphere.

Data collection and processing. Data sets were measured on Fuji imaging plates at the X-4A beamline at the NSLS, Brookhaven National Laboratory. Crystals were stabilized and cryoprotected before cooling to -160 °C by soaking in a solution of 25% PEG 8000, 200 mM MgSO₄, 100 mM HEPES, pH 7.0 and 22% xylitol. Two crystal forms of the complex were examined: one (native) form has space group *P*6₁22, with $a = b = 91.1$ Å, two aFGF protomers and one decasaccharide chain in the asymmetric unit; the other (selenomethionyl) form has space group *C*222₁, with $a = 87.7$ Å, $b = 158.1$ Å, $c = 190.6$ Å, six aFGF protomers and three decasaccharide chains in the asymmetric unit. One single-crystal data set was collected on each crystal form. MAD data were collected at four wavelengths around the K selenium edge. Imaging plates were exposed and digitized using a Fuji scanner. Raw data were converted to integrated intensities by DENZO²¹. Reflections were scaled and merged with the CCP4 programs²² (native) or left unmerged (MAD).

Phases for the native aFGF complex were obtained by molecular replacement with search model 1AFC (bovine aFGF)⁵ using AMoRe²³. Molecular replacement phases for the semet-aFGF complex were obtained by AMoRe using the native aFGF dimer model. Experimental phases and figures of merit for the semet-aFGF complex were obtained using MADSYS programs²⁴. Molecular replacement and experimental phases were combined by ABCDCOMB (W.A.H.). In both crystal lattices, the aFGF molecules are organized in repeated motifs, namely, a tetrameric protein complex in which heparin-linked aFGF dimers are associated laterally across crystallographic two-fold axes by protein-protein interfaces (519 Å² per protomer). These lattice interactions are not likely to be biologically relevant as residues involved are not conserved among FGF subtypes and our site-directed mutagenesis of Ala 103, central in this interface, does not affect aFGF efficacy.

Model building and refinement. Refinement was done and maps were calculated using XPLOR²⁵. Maps were viewed with program O (ref. 26). For semet-aFGF (*C*222₁), after partial refinement of the initial protein model, sigma-weighted electron-density maps were used to position hexose units of heparin starting from NMR structure 1HPN⁸. Refinement was carried out using non-crystallographic restraints in all but the final cycles of refinement. A solvent mask was calculated and a correction for bulk solvent was applied to FPART. All aFGF side chains were built in for each protomer (PDB segment

identifications (SEGIDs) A1, A2, B3, B4, C5 and C6) and five hexose units in each of the three decasaccharide chains (SEGIDs: ASUG, A_{Sugar}; BSUG, B_{Sugar}; CSUG, C_{Sugar}) were located. The model was refined using solvent-corrected data between 13–3 Å Bragg spacings. For native aFGF (*P*6₁22), the semet-aFGF dimer was used as the initial model refined against the native data. Structure refinement was carried out as above. Six monosaccharide units of the single heparin decasaccharide chain (SEGID: HSUG, H_{Sugar}) and all side chains of the aFGFs (SEGIDs: H7 and H8) were built into the maps. The model was refined for data between 13–2.9 Å Bragg spacings. Both structures exhibit good geometry and have no residues in the disallowed regions of the Ramachandran plot²⁷.

Received 16 February; accepted 15 April 1998.

1. Basilico, C. & Moscatelli, D. The FGF family of growth factors and oncogenes. *Adv. Cancer Res.* **59**, 115–165 (1992).
2. Jaye, M., Schlessinger, J. & Dionne, C. A. Fibroblast growth factor receptor tyrosine kinases: molecular analysis and signal transduction. *Biochem. Biophys. Acta* **1135**, 185–199 (1992).
3. Schlessinger, J. Signal transduction by allosteric receptor oligomerization. *Trends Biochem. Sci.* **13**, 443–447 (1988).
4. Spivak-Kroizman, T. et al. Heparin-induced oligomerization of FGF molecules is responsible for FGF receptor dimerization, activation, and cell proliferation. *Cell* **79**, 1015–1024 (1994).
5. Zhu, X., Hsu, B. T. & Rees, D. C. Structural studies of the binding of the anti-ulcer drug sucrose octasulfate to acidic fibroblast growth factor. *Structure* **1**, 27–34 (1993).
6. Ornitz, D. M. et al. FGF binding and FGF receptor activation by synthetic heparan-derived di- and trisaccharides. *Science* **268**, 432–436 (1995).
7. Faham, S., Hileman, R. E., Fromm, J. R., Linhardt, R. J. & Rees, D. C. Heparin structure and interactions with basic fibroblast growth factor. *Science* **271**, 1116–1120 (1996).
8. Mulloy, B., Forster, M. J., Jones, C. & Davies, D. B. NMR and molecular-modelling studies of the solution conformation of heparin. *Biochem. J.* **293**, 849–858 (1993).
9. Burgess, W. H., Shaheen, A. M., Hampton, B., Donohue, P. J. & Winkles, J. A. Structure-function studies of heparin-binding (acidic fibroblast) growth factor-1 using site-directed mutagenesis. *J. Cell Biol.* **45**, 131–138 (1991).
10. Thompson, L. D., Pantoliano, M. W. & Springer, B. A. Energetic characterization of the basic fibroblast growth factor-heparin interaction: identification of the heparin binding domain. *Biochemistry* **33**, 3831–3840 (1994).
11. Ornitz, D. M., et al. Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells. *Cell Biol.* **12**, 240–247 (1992).
12. Moy, F. J. et al. Properly oriented heparin-decasaccharide-induced dimers are the biologically active form of basic fibroblast growth factor. *Biochemistry* **36**, 4782–4791 (1997).
13. Herr, A. B., Ornitz, D. M., Sasisekharan, R., Venkataraman, G. & Waksman, G. Heparin-induced self-association of fibroblast growth factor-2: evidence for two oligomerization processes. *J. Biol. Chem.* **272**, 16382–16389 (1997).
14. Springer, B. A. et al. Identification and concerted function of two receptor binding surfaces on basic fibroblast growth factor required for mitogenesis. *J. Biol. Chem.* **269**, 26879–26884 (1994).
15. Kan, M. et al. An essential heparin-binding domain in the fibroblast growth factor receptor kinase. *Science* **259**, 1918–1921 (1993).
16. Weismann, C. et al. Crystal structure at 1.7 Å resolution of VEGF in complex with domain 2 of the Flt-1 receptor. *Cell* **91**, 695–704 (1997).
17. van Holde, K. E. & Weischet, W. O. Boundary analysis of sedimentation-velocity experiments with monodisperse and paucidisperse solutes. *Biopolymers* **17**, 1387–1403 (1978).
18. Huang, J., Mohammadi, M., Rodrigues, G. A. & Schlessinger, J. Reduced activation of RAF-1 and MAP kinase by a fibroblast growth factor receptor mutant deficient in stimulation of phosphatidylinositol hydrolysis. *J. Biol. Chem.* **270**, 5065–5072 (1995).
19. Jaye, M. et al. Human endothelial cell growth factor: cloning, nucleotide sequence, and chromosome localization. *Science* **233**, 541–545 (1986).
20. Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **9**, 1665–1672 (1990).
21. Otwinowski, Z. DENZO 1-56-62 (SERC Daresbury Laboratory, Warrington, UK, 1993).
22. Collaborative Computational Project no. 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
23. Navaza, J. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
24. Hendrickson, W. A. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**, 51–58 (1991).
25. Brünger, A. T. *X-PLOR Version 3.1, A System for X-ray Crystallography and NMR* (Yale University Press, New Haven and London, 1992).
26. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
27. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
28. Evans, S. V. SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graph.* **11**, 134–138 (1993).
29. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).

Acknowledgements. We thank C. Bingham, M. Cuff, L. Shapiro, H. Yamaguchi and C. Ogata for help in data collection; J. Arnez, K. Drickamer, B. Honig, S. Hubbard, J. Passner, M. Quesenberry, L. Shapiro, and H. Yamaguchi for discussions; P. Østergaard (Novo Nordisk) for heparin oligosaccharides; M. Gawinowicz (HHMI Columbia) for amino-acid analysis; D. King for electrospray mass spectrometry; and C. Turgeon and J. Hansen for analytical ultracentrifugation. This work was supported in part by a grant from the NIH. Beamline X4A at the National Synchrotron Light Source, a DOE facility, is supported by the HHMI.

Correspondence and requests for materials should be addressed to W.A.H. (e-mail: hendw@convex.hhmi.columbia.edu). Coordinates are deposited in the Protein Data Bank with accession numbers 1AXM and 2AXM.

Gene research products

Tools for genetic research covered in this feature include a hand-held laboratory 'assistant', high-grade nucleotides for PCR, a probe for murine telomeres, a catalogue for antisense oligonucleotides and new gel loading pipette tips.

CEQ 2000

From Beckman Coulter

Capillary electrophoresis DNA sequencing in a 96-well format

Utilizing capillary electrophoresis (CE) technology with laser-induced fluorescence detection, this system features four-colour, dideoxy-terminator chemistry and operates in a microplate format. CEQ handles more of the activities of dideoxy-terminator sequencing, from pre-run setup, sample denaturation, gel replenishment, injection and separation, to final determination of base sequence, with little or no human intervention. Using four dyes per capillary to help deliver fast turnaround and high throughput, a final report calling over 500 bases on eight samples can be produced in just 2 h, and on a complete 96-well plate in 24 h. In addition, the system's software measures the accuracy of each base call, eliminating the need to access a separate software program for accuracy determination. CEQ samples have format compatibility with the Biomek 2000 laboratory automation workstation and the Multimek 96 automated 96 channel pipettor.

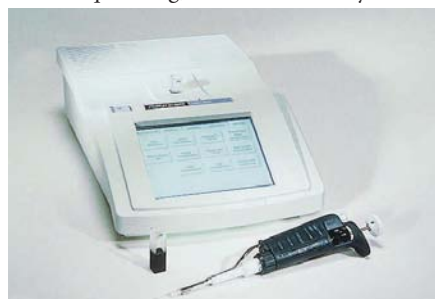
Reader Enquiry No. 100

MBA 2000 system

From Perkin-Elmer

A new productivity tool for genetic analysis, biotechnology and drug discovery

The molecular biology assistant (MBA) instrument is a compact computational system that consolidates and organizes routine and customized functions used in virtually all protocols that employ biochemicals. The system employs an intuitive, touchscreen display and has no moving parts, all in a nine-inch wide, low-profile package. Choices are made simply with a touch-to-select action and the user is prompted as needed. From estimating reagent usage costs to checking absorbance, from setting up serial dilutions and PCR master mixes to designing new oligonucleotides for PCR or DNA sequencing, the MBA 2000 system is



Getting an MBA from Perkin-Elmer may pay off.

designed to make computations and quantification accessible and rapid. It can be customized for specific analyses and determinations.

Reader Enquiry No. 101

Platinum PCR products

From Gibco BRL

New literature covering the Platinum Taq polymerase and its corresponding antibody

These products provide automatic hotstarts for greater specificity in PCR procedures. The polymerase-antibody mixture is said to eliminate the need to add antibody before assembling PCR. Provided in Taq equivalent units, the antibody is supplied as a solution to be added to PCR by the user. There is said to be real-time activation of Platinum Taq DNA polymerase with this system, and the Taq antibody is said to eliminate pre-PCR misprimings. Automatic hotstarts activate Taq at 94 °C during the first step of PCR denaturation. The product design is said to ensure stable storage by completely inhibiting the enzyme until use. According to the manufacturer, reactions can be initiated at room temperature, and amplify templates up to 5 kb.

Reader Enquiry No. 102

PCR-grade nucleotides

From Boehringer Mannheim

For applications that require high-quality reagents and provide quality results

These deoxynucleotides have been HPLC purified to high standards (HPLC area ≥ 99 percent, dNDP area ≥ 0.9 per cent) and are supplied as ready-to-use sodium salt solutions. Each lot is function tested in an RT-PCR assay to confirm the absence of RNase, DNase and nicking activity. These reagents should prove useful in reverse transcription (RT), PCR, RT-PCR and cycle sequencing. The nucleotides (dATP, dCTP, dGTP, dTTP and dUTP) can be purchased in 25- μ mol and 1,250- μ mol sizes.

Reader Enquiry No. 103

T3 thermocycler

From Biometra

Part of a new generation of multiblock instruments

These instruments feature three independently operating sample blocks within one compact housing. Each block is equipped with its own pressure-optimized, heated lid. This can be set to run its own individual programme, allowing the user to run three different protocols at any one time. The



Biometra's T3 three-in-one thermal cycler.

triple format can also significantly increase the capacity of a laboratory by enabling three users to work simultaneously on one instrument. The system also has the ability to optimize run conditions quickly.

Reader Enquiry No. 104

Blotting materials

Hybond-XL

From Amersham Pharmacia Biotech

A new positively charged nylon membrane for radioactive nucleic acid blotting

This is the company's first membrane designed specifically for radioactive nucleic acid blotting. It is built to provide an improved signal-to-noise ratio in applications such as Southern blots, northern blots and colony/plaque lifts. This means that shorter exposure times can be used and that more information can be obtained. It is also suitable for re-probing. Moreover, improvements in signal can be obtained if protocols are optimized for the new membrane, which is said to be well suited for use with such products as Rediprime, Megaprime and 32 P for Southern blotting applications. Hybond-XL is quality controlled to ensure detection of 0.1 pg of genomic DNA in a Southern blot with a four-hour exposure.

Reader Enquiry No. 105

Protran BA 83 and 85

From Schleicher & Schuell

Nitrocellulose transfer membranes for blotting techniques in life-science applications

Protran BA 83 (0.2 μ m) and BA 85 (0.45 μ m) are said to be versatile membranes that are well suited to Southern, northern and western blotting, diagnostic antigen

detection and colony/plaque lifts. The membrane should have excellent handling properties as a result of the very high mechanical stability of the nitrocellulose material. It has an increased tensile strength, stated to be double that of previous membranes, together with low background noise. The high mechanical stability of this pure nitrocellulose membrane gives improved handling characteristics with reduced brittleness.

Reader Enquiry No. 106

An RNA world

RNApure

From GenHunter

For the extraction of total RNA from tissues or cells

GenHunter has released its new RNApure reagent — a simple monophasic solution for rapid, one-step isolation of intact total RNA. According to the manufacturer, the isolated RNA is of high quality and can be used for differential display, northern or reverse northern-blot analysis. Both 50-ml and 100-ml containers are available, and come with detailed protocol.

Reader Enquiry No. 107

RNA labelling systems

From Enzo Diagnostics

Newly formatted RNA transcript labelling

These systems have been formulated and optimized for use with microchip assays, gene expression analyses, RNase protection assays and other non-radioactive applications. RNA probes can be labelled for use in *in situ* hybridization assays, Southern and northern blot analyses, and other hybridization procedures. Reagents and a detailed protocol are provided with these systems, and each provides a reagent kit containing enzymes (T3/T7, T7/SP6, T3, T7 or SP6 RNA polymerases), buffers and control templates. There is also a choice of nucleotide labelling mix, which includes biotin-11-UTP, biotin-16-UTP, biotin-11-CTP, biotin-17-ATP, digoxigenin-11-UTP or fluorescein-12-UTP.

Reader Enquiry No. 108

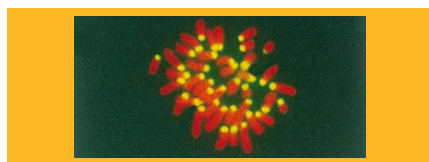
DNA tools

Pan centromeric DNA probe

From Oncor

An addition to the company's line of mouse paint and unique sequence DNA probes

Intended for use in fluorescence *in situ* hybridization (FISH), this new biotin-labelled probe targets the centromeres of all twenty-one mouse chromosomes. According to the manufacturer, it takes just under three hours for mouse bone marrow assayed with the probe to yield a large, bright fluorescent spot or signal at the site of each centromere, with a clean background. The mouse



FISHing for a mouse centromere with Oncor.

pan-centromeric DNA probe is biotin-labelled and supplied pre-mixed in hybridization buffer in a quantity of 500 µl. This amount is sufficient to assay 20–50 slides, depending on hybridization area.

Reader Enquiry No. 109

cDNA libraries

From Stratagene

The reagents needed for constructing a cDNA library

Stratagene has optimized its cDNA synthesis kit to provide full-length, high-quality cDNA to researchers. Combined with cloning vectors, high-efficiency lambda packaging extracts and specialized plasmid transformation technology, these cDNA kits can provide users with a large, complex cDNA library. Optimization of the synthesis process includes the first-strand synthesis, the directional cloning method and even the polishing reaction that is used to create blunt ends for efficient adapter ligation. Efficient synthesis from mRNA to cDNA is performed using a recombinant Moloney murine leukaemia virus reverse transcriptase (MMLV-RT) — employed for its ability to produce large quantities of high-molecular-weight cDNA.

Reader Enquiry No. 110

1998 Biognostik catalogue

From Chemicon

A range of antisense oligonucleotides capable of selective inhibition of gene expression

Biognostik offers a large selection of antisense oligonucleotides, with brand names such as PCR PrimeR and HybriProbes. These products should prove useful in a wide variety of biological research areas. Custom oligonucleotide kits designed against a specific target gene are also available.

Reader Enquiry No. 111

MPG streptavidin and avidin

From Quantum Biotechnologies

Magnetic porous glass particles with streptavidin or avidin covalently bound

This system is designed to allow for fast, reliable magnetic separation of biotinylated molecules. The biotin binding capacity of MPG streptavidin and avidin ($K_d=10-15$ M) is retained under a variety of conditions, including treatment with solvents (ethanol and DMSO, for example) and detergents (such as, Triton X-100 and Tween 20). Routine applications include solid-phase

PIERCE

SuperSignal® ULTRA Chemiluminescent Substrate*

Femtomogram Sensitivity for HRP Detection in Western Blotting

Features/Benefits:

- Mid-femtomogram-level (10^{-15}) sensitivity at an economical price
- 24-hour light emission allows for multiple exposures
- Recommended primary antibody dilution is 1:5,000- 1:50,000; recommended secondary antibody dilution is 1:50,000-1:250,000
- 24-hour working solution stability
- Allows for stable hard-copy results on film, rapid processing and high resolution

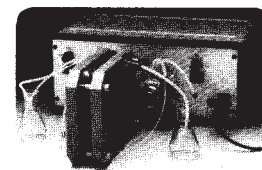
Tel: 800-874-3723 or 815-968-0747 • Fax: 815-968-7316
CS@piercenet.com • Internet: www.piercenet.com

* Patent Pending

READER ENQUIRY NO. 34

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5–10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

P.O. Box 650126, D-81215 Munich
For USA call 1-800 Liposome

READER ENQUIRY NO. 62

Check out
our homepage at
<http://www.resgen.com>

Peptides
\$18/residue

Antibodies
monoclonal - \$8,500
polyclonal - \$1,195

Complete services for your custom peptide and polyclonal or monoclonal antibody needs.

Research Genetics, Inc.
U.S. or Canada 800-533-4363
U.K. 0-800-89-1393
FAX 256-536-9016

READER ENQUIRY NO. 37

nucleic acid sequencing, nucleic acid probe preparation and purification, the isolation and enrichment of cDNAs, single-strand DNA labelling and library construction, and specific hybridization. MPG LCA (long-chain alkylamine) is used for ligand attachment. It is derivitized to provide amino groups that offer greater flexibility for binding a wide variety of ligands. The long, 15Å spacer arm present on the surface of the particle enhances reactivity by minimizing steric hindrance. MPG LCA beads can be used for the purification of nucleic acids, cDNA amplifications and subtractive hybridization. **Reader Enquiry No. 112**

Making things gel

AlphaArchiver 600

From Alpha Innotech

Designed as an entry-level, fully upgradeable, gel documentation system

According to the manufacturer, this system provides documentation and archiving of gels, films and membranes in just minutes. It utilizes CCD-camera technology and optics to capture real-time (30-fps) images, at 762 × 494 (V × H) pixel resolution and greater than 56 db signal-to-noise ratio. It can also perform low-light, chemiluminescence applications with the addition of the DE-300 EconoImage light cabinet or the DE-400 MultiImage light cabinet. It is also available with several printing options for hard copy results in seconds rather than minutes. A fully integrated software package is also provided for image documentation and archiving in most current file formats. **Reader Enquiry No. 113**

EZ Load molecular rulers

From Bio-Rad

Precision-sized DNA markers, blended with loading buffer for ease and convenience

These rulers may be loaded directly onto the electrophoresis gel. They are available in six size ranges: 20–1,000 bp, 10–1,000 bp, 100–3,000 bp, 500–8,000 bp and 1–15 kb, as well as a precision molecular mass standard. According to the manufacturer, there is no spurious migration from sequence variation and no contaminating plasmid DNA. Visually distinct reference bands simplify readability, and the rulers progress by precise, easily calculable increments for easy size determination, says Bio-Rad. **Reader Enquiry No. 114**

DNA ladders

From Flowgen

Four DNA ladders for DNA electrophoresis

These new ladders now cover molecular sizes from 20 bp to 8 kb in steps of 20, 100 or 500 bp, and are said to be free from contaminating plasmid or 'junk' DNA. The first

of the four available ladders is a 20-bp increment DNA ladder, consisting of 50 fragments in the range from 20 to 1,000 bp. The 100-bp ladder consists of ten fragments from 100–1,000 bp in 100-bp increments. There is also a 100-bp, extended-range ladder with over 30 fragments covering the range of 100 bp to 3,000 bp in 100-bp increments. The last member of the group is a 500-bp version with 16 fragments from 500 bp to 8 kb, occurring at 500-bp increments. **Reader Enquiry No. 115**

Last but not least

S120H oven/shaker

From Merck

A hybridization oven/shaker from Stuart Scientific

This unit provides a value alternative for those undertaking hybridizations or general shaking/incubator procedures. Supplied by Merck throughout Europe, the S120H features high-performance temperature control, using forced air circulation and a variable-speed, seven-bottle rotisserie. It has been designed for the widely used 20-cm wide hybridization membrane. The door is counterbalanced to allow for easy opening and designed to ensure full front and top access. The rotisserie is easily removed and serves also as a bottle stand. **Reader Enquiry No. 116**



Hot tips — SequaStrip tips from Embi Tec.

SequaStrip tips

From Embi Tec

A system that is touted to allow the loading of sequencing gels in as little as five minutes

According to the manufacturer, this is the only eight-tip strip and multichannel pipette combination designed specifically for DNA sequencing. The tips are short with a very flat, narrow paddle and the strip is straight, which is said to result in a simple loading action. For a precise draw each time, the pipette is calibrated. The strips were designed for sequencing systems and fit 0.2-, 0.3- and 0.4-mm gels. The tips are disposable, to help eliminate broken glass or cross-contamination. **Reader Enquiry No. 117**

These notes are compiled by Brendan Horton from information provided by the manufacturers.

For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

**Detect residual DNA with Wako's
DNA Extractor Kit
or Isolate Genomic DNA with
DNA Extractor WB Kit**
Both use the Nal Method
No Phenol No Chloroform

Wako
Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 35

**Greenland
Summit
Ice Cores
CD-ROM**

Cores from the Greenland Summit provide information on environmental change. Data measurements include major ions, isotopes, sulfate, nitrate, CO₂, ecm, etc. To order a CD-ROM contact:

nsidc@kryos.colorado.edu
Phone +1 (303) 492 6199
Fax +1 (303) 492 2468

Data are available electronically at:
<http://arcss.colorado.edu>

READER ENQUIRY NO. 36

New Severn Super Sequencer Acrylamide Gel Solution

Severn Super Sequencer is a new improved gel solution for use in automated DNA sequencing. Complete with denaturant and buffer, it comes in a series of 'ready to use' (RTU) formulations. Simply pour and Polymerise.

So in the race to sequence, don't walk, *RUN!*

With *Severn Super Sequencer* – Available in 100ml, 500ml and 5000ml volumes.

Trial 100ml available on request.

Severn Biotech Ltd

Severn Biotech Ltd, Unit 2, Park Lane, Stourport Road, Kidderminster, Worcs DY11 6TJ, U.K.
Tel: (01562) 825286 Fax: (01562) 825284 E-Mail: 101626.2102@Compuserve.com

UK Manufacturers of Liquid Acrylamides, Buffers and Electrophoretic Media

READER ENQUIRY NO. 37